THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY

NOTES FOR AUTHORS

From the next issue the Editor will be Professor R. Davis of the Department of Psychology, University of Newcastle, Claremont Place, Newcastle upon Tyne. All manuscripts should be submitted to him. Only original papers are accepted for publication. One copy must be a *typewritten original* with double spacing and wide margins. Figures should be drawn to professional standards in Indian ink on tracing linen or Bristol Board; the original drawing (plus one copy) should be submitted. Allowance must be made for a reduction in size when the drawing is reproduced (the width of lines as reproduced should be between 0·25 and 0·33 mm). Explanations must be placed in the figure legend; the drawing should contain as little lettering as possible and lettering should be inserted in *blue* pencil or on a tracing paper to fit over the drawing. Tables and figure legends should be typed on separate sheets of paper. Where possible related diagrams should be grouped together to form a single figure.

Intending contributors should pay attention to the *Journal*'s conventions, and in particular to the form of headings, and sub-headings used, and references (in the Reference list titles of journals will in future be printed in full without abbreviation). The *Journal* will use the *Système Internationale* for all measurements (if terms like inches are used, metric equivalents must be quoted). The use of standard abbreviations in the text is permissible (if in doubt, write term in full the first time it is used and give abbreviation in brackets); please note that the *Journal* does not use S for subject.

An abstract must be included. Although abstracts should be as tersely written as possible, they should summarize all the main experimental findings. The length of abstracts will vary with the number and complexity of the experiments reported but should usually be between 50 and 400 words.

An initial decision on a paper will normally be reached within 3 months of receipt. Where only minor revisions are requested, date of publication will be determined by date of receipt of the paper provided a satisfactory revision is received within 1 month of the request for a revision being made. The delay between final acceptance of a paper and its publication will normally be between 3 and 6 months. The date on which the manuscript is received in its final version (apart from minor alterations—see above) is printed.

Contributors receive 25 copies of their paper free. If they require further copies they should apply to the Publishers at the time of returning proofs.

CONTRAST AND SIZE VARIABLES AND THE TILT AFTER-EFFECT

D. M. PARKER

Department of Psychology, University of Durham

The relationship between the contrast and bar width of adapting and test gratings in the determination of the magnitude of the tilt after-effect was explored. When the contrast of adapting and test gratings is varied concomitantly over four levels, the magnitude of the tilt after-effect does not change. When the contrast of the adapting grating is higher than the test grating, the magnitude of the after-effect is increased and when its contrast is lower than the test grating, the magnitude of the after-effect is decreased. When the bar width of the test grating is narrower or broader than that of the adapting grating, the magnitude of the after-effect does not change. The implications of these results for a neurophysiological explanation of the tilt after-effect are discussed.

Introduction

Gibson and Radner (1937) reported that if an observer gazes for a period of time at a line that is tilted from the vertical, when he subsequently looks at a vertical line, it will appear tilted away from the previously inspected line, i.e. the observer must move the truly vertical line in the direction of the tilted adapting line in order that it should appear vertical. They named this the "tilt after-effect" and explained it by supposing that an observer had at his disposal a system of oppositional scales, one of them being tilt-right versus tilt-left, and that the neutral point on this scale (the vertical) "normalized" in the direction of the tilted line during the period of adaptation. Since the original report there have been a number of attempts to subsume this phenomenon under a more general mechanism of figural after-effects, the most notable being those espoused by Kohler and Wallach (1944) and Osgood and Heyer (1952). Both of these theories have been extensively criticized and it is clear that some other explanation of the tilt after-effect is demanded (McEwen, 1958; Sutherland, 1961; Over, 1971). Sutherland (1961) proposed that if the human visual system contains neurones similar to those found in the striate cortex of the cat (Hubel and Wiesel, 1959), a ready explanation of the tilt after-effect is available. If a line of a particular orientation is exposed to an observer, it will result in a high level of activity in those neurones maximally sensitive to that orientation and some activity in those most responsive to neighbouring orientations. If as a result of prolonged activity in these neurones adaptation occurs, then exposure of a line at a slightly different orientation in the same retinal region will result in maximal activity in those neurones which are normally maximally sensitive to contours whose orientations lie further away from the adapting line. If the orientation at which a line is seen depends on the ratios of activity in a neighbouring group of orientation detectors then the observer will see the test line as being tilted

1

further from the adapting line than in fact it is. This type of explanation has been more fully developed in a recent paper by Coltheart (1971).

If this explanation is correct, then one might reasonably suppose that variables that are known to affect the activity of visual cortex neurones in cat and monkey would also influence the tilt after-effect. One variable which clearly affects the firing rates of these neurones is the contrast of the stimulus (Campbell, Cooper and Enroth-Cugell, 1969a; Campbell, Cooper, Robson and Sachs, 1969b). A second variable is the spatial frequency of the stimulus (bar width). In both experiments it was found that not only are visual cortex units responsive to the contrast and orientation of the visual stimulus but also to its spatial frequency, so that a given unit is not appreciably affected by a grating whose spatial frequency lies one octave above or below its optimum frequency. Psychophysical evidence has been presented that suggests that similar spatial selectivity may be present in the human visual system (Campbell and Robson, 1968; Blakemore and Campbell, 1969). If the tilt after-effect can be explained by assuming that adaptation occurs in neurones in the human visual cortex similar in type to those reported for cat and monkey, then it should be possible to alter the magnitude of the after-effect by manipulating the relationship between the contrast and spatial frequency of adapting and test stimuli. This report describes experiments which explore the relationship between these variables and the magnitude of the tilt after-effect.

Method

Subjects

Subjects were all members of the Psychology Department at Durham. Apart from one subject (shown as ⟦○——○⟧ in all figures) who did all four experiments, each curve represents data from one subject. Nine subjects were used in all.

Apparatus

A vertical grating pattern, with a sinusoidal intensity distribution was displayed on an oscilloscope screen (Tektronix 561A with P31 phosphor) using the method of Schade (1956) as modified by Campbell and Green (1965). The sine wave voltage applied to the beam brightness control modulated the brightness of the grating pattern about a mean level and did not affect the space average luminance of the screen. The contrast of the grating pattern is defined as $I_{max} - I_{min} / I_{max} + I_{min}$ (I_{max} = luminance of brightest part of the grating, I_{min} = luminance of dimmest part of the grating). Contrast was varied by placing a logarithmic attenuator (Advance type A. 64) between the sine wave generator (Aim TFO 128) and the Z–axis input of the oscilloscope. Because the luminance of the beam brightness control is a non-linear function of the voltage applied to it the actual contrast of the gratings used in this study was determined independently of the voltage. Low frequency square wave gratings, produced by driving the beam brightness control with square waves of the same peak to peak amplitude as the higher frequency sine wave used as the stimulus in the present investigation, were displayed on the oscilloscope screen and the luminance of the brightest and dimmest parts measured with an S.E.I. photometer. The contrast values quoted are the means of readings taken by two observers. The oscilloscope beam brightness control was driven at 85, 27, 8·5 and 2·7 V, peak to peak, and the corresponding contrast values obtained were 0·95, 0·9, 0·32 and 0·085 respectively. The space average luminance of the oscilloscope screen remained constant at approximatley 6·0 cd/m².

A mask with a circular aperture was placed in front of the oscilloscope screen and a small black circle was drawn on the screen itself. The subject's head was supported by a chin rest and he looked through a small aperture, behind which was a Dove prism. The stimulus

field was located at infinity by use of a suitable correcting lens. By moving a lever the subject could rotate the Dove prism and thus change the orientation of the grating. The stimulus field subtended 3·24° and the fixation circle at its centre subtended 0·7°. The spatial frequency of the sinusoidal grating was 4·6 cycles/degree, except in Experiment IV where the oscilloscope time base was switched to double and half the spatial frequency of the adapting stimulus during resetting. Viewing conditions were monocular and a system of baffles ensured that nothing was visible to the subject but the stimulus field.

Procedure

The subjects' mean estimate of the vertical was first obtained in the absence of any pre-adaptation by getting him to reset the grating to the vertical, after it had been tilted, by moving the lever. When the subject adapted to the tilted grating he was instructed to fixate the perimeter of the small central circle and to keep his eye moving around its circumference in order to avoid formation of an after image. In all experimental conditions the subject was adapted for 50 sec to a grating pattern tilted 10° from the vertical in an anti-clockwise direction. This stimulus will subsequently be called the adapting stimulus. Immediately after this period of adaptation, the contrast of the grating was either increased, decreased or allowed to remain the same, and the subject was asked to reset the grating to the vertical by moving the lever. This stimulus will subsequently be called the test stimulus. Subjects were instructed to move the grating past their estimated setting and return to it when re-estimat-ing the vertical following a period of adaptation in order to avoid a tendency to stop short of their true estimate when moving in towards it (Gibson and Radner, 1937). The subject was allowed a minimum of 2 min recovery between each adapting and resetting trial and checks were run during the experimental session to ensure that the unadapted estimate of the verti-cal did not change appreciably.

In Experiment I the contrast of the adapting and test grating remained the same and the extent of the tilt after-effect of four contrast levels was determined. In Experiment II the contrast of the test grating was kept constant at 0·95; the extent of the after-effect was determined at four levels of adapting contrast, one equal to, and three lower than the contrast of the test grating. In Experiment III the contrast of the test grating was held constant at 0·085 and the extent of the after-effect was again determined for four levels of adapting contrast, one equal to, and three higher than the contrast of the test grating. In Experiment IV, following adaptation to the high contrast 4·6 cycles/degree adapting stimulus, the contrast remained the same but the spatial frequency of the test stimulus was either doubled or halved before the subject reset the grating to the subjective vertical.

Results

Figure 1 shows the results of Experiment 1 where the contrast of adapting and test gratings were varied over four levels of contrast. The points plotted in this and subsequent figures are the mean pre-adaptation settings minus the mean post adaptation settings. Each curve shows the results of one subject. It is clear that there is no systematic effect of changing the absolute contrast of adapting and test gratings on the magnitude of the tilt after-effect.

The results of Experiment II are shown in the left hand-graph in Figure 2. It can be seen that as the contrast of the adapting grating is lowered relative to the test grating contrast, the magnitude of the after-effect decreases from a mean value of 3·43° for the three subjects when adapting and test contrasts are equal to a mean value of 0·43° when the adapting contrast is rather more than 1 log. unit below the test contrast. An analysis of variance performed on these data shows that the contrast effect ($F = 17.57$; $df = 3,6$; $P < 0.01$) and the differences between subjects ($F = 19.625$; $df = 2,6$; $P < 0.01$) are significant. The mean squares for

the subject and contrast effects were tested against the Subjects × Contrast interaction, which was also significant ($F = 2\cdot85$; $df = 6,48$; $P < 0\cdot05$). Because the contrast levels were not equally spaced, orthogonal coefficients for the unequal
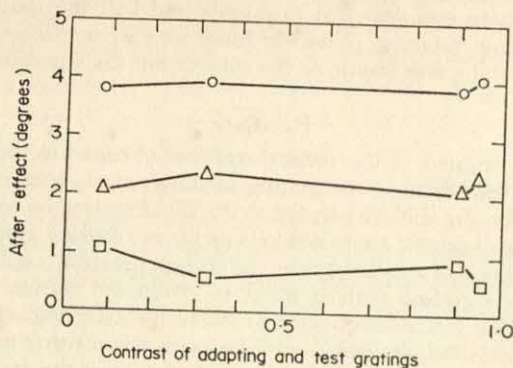

FIGURE 1.   The magnitude of the tilt after-effect for three subjects when the contrast of adapting and test gratings were varied concomitantly over four contrast levels.

intervals used, were calculated using the method of Grandage (1958) in order to perform the trend analysis in this and the subsequent experiment. The linear regression component was highly significant ($F = 76\cdot99$; $df = 1,6$; $P < 0\cdot001$) and the deviations from linear trend were not significant ($F = 2\cdot413$; $df = 3,6$; $P > 0\cdot05$).
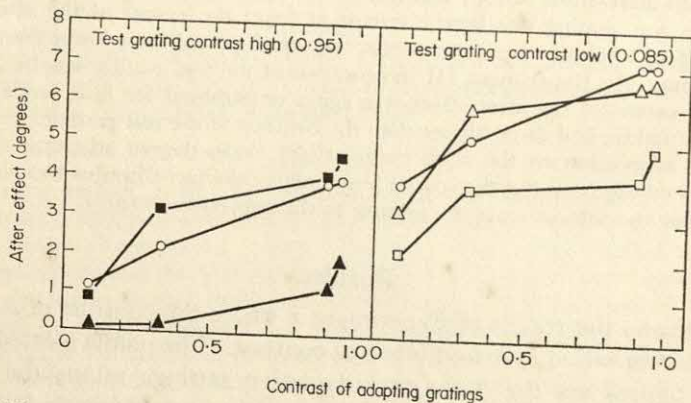

FIGURE 2.   The left-hand graph shows the magnitude of the tilt after-effect for three subjects when the contrast of the adapting grating is lower than or equal to the test grating. The right-hand graph shows the magnitude of the after-effect when the contrast of the adapting grating is equal to, or higher than the test grating.

The results of Experiment III are plotted on the right-hand graph in Figure 2. It can be seen that as the contrast of the adapting grating is increased relative to the test grating contrast the after-effect increases from a mean of $3\cdot13°$ where the adapting and test contrasts were equal, to a mean value of $6\cdot2°$ when the adapting contrast is rather more than 1 log. unit higher than the test contrast. The analysis of variance shows a significant effect of contrast levels ($F = 48\cdot72$; $df = 3,6$; $P < 0\cdot001$) and a significant difference between subjects ($F = 48\cdot282$; $df = 2,6$;

$P < 0.001$). In this experiment the Subjects × Contrast interaction was not significant ($F = 0.87$; $df = 6,48$; $P > 0.05$). The trend analysis showed that both linear ($F = 112.72$; $df = 1,6$; $P < 0.001$) and quadratic ($F = 26.06$; $df = 1,6$; $P < 0.01$) regression components were significant. Deviations from first and second order trends were insignificant.
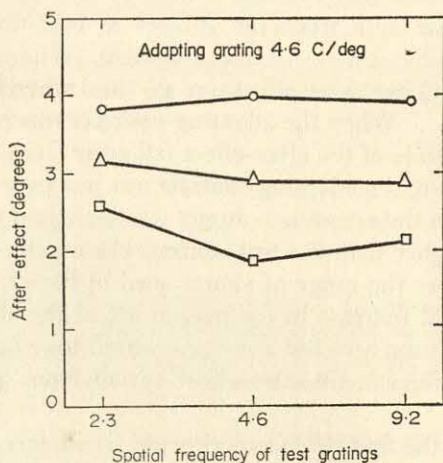

FIGURE 3. The magnitude of the tilt after-effect for three subjects when the spatial frequency (bar width) of the test grating is one octave below, the same or one octave higher than that of the adapting grating.

Figure 3 shows the results of Experiment IV, where following adaptation to a grating of 4·6 cycles per degree, the spatial frequency (bar width) of the test grating was either the same, lower or higher than the spatial frequency of the adapting grating. It can be seen that there is no systematic effect of changing the spatial frequency of the test stimulus on the magnitude of the tilt after-effect.

Discussion

The finding that absolute variations in the contrast of adapting and test gratings does not affect the magnitude of the tilt after-effect is not too surprising. If the activity of visual cortex units is monotonically related to the contrast of the stimulus and the orientation at which a line is seen depends on some measure like the weighted average of activity in a group of neighbouring orientation detectors, then such a result would be expected. Day (1962) reported that the magnitude of this after-effect is not modified by wide variations in "luminance contrast". Unfortunately, he did not specify whether he was referring to absolute or relative contrast. From the results of the present experiment it would appear that he was referring to conditions where the adapting and test contrasts were varied in unison.

The results of Experiments II and III are quite clear. They show that if the contrast of the adapting grating is lower than that of the test grating the magnitude of the after-effect is decreased and when it is higher it is increased. These findings are clearly in accord with the explanation of this effect proposed by Sutherland

(1961). Thus exposure of a low contrast grating will result in less firing in the "orientation detectors" responsive to that and neighbouring orientations than if a high contrast stimulus is used. Consequently, there will be less adaptation and when a high contrast grating is exposed at a neighbouring orientation it will be less affected and the apparent shift in orientation will be less. It should be noted that Experiment I is complementary to Experiments II and III. This can be appreciated by examining the data from the subject in common between the three Experiments (indicated by ○ —— ○ in the figures). When the adapting and test contrast were high (0·95) her mean affect was 3·9° and when both were low (0·085) her mean affect was 3·8°. When the adapting contrast was reduced relative to the test contrast the magnitude of the after-effect fell away from this level and reached a mean level of 1·2° where the adapting contrast was just over 1 log. unit lower than the test contrast. When the adapting contrast was increased to a level where it was just over 1 log. unit higher than the test contrast the magnitude of her after-effect reached 7·2°. Thus over the range of values used in Experiments II and III this subject showed a sixfold increase in the magnitude of the after-effect. Where, in Experiment I, the adapting and test contrasts varied together over four levels the mean extent of this subject's tilt after-effect varied from 3·8° to 4·0° in an unsystematic way.

While the results of the first three experiments are in agreement with the model proposed by Sutherland (1961) the results of Experiment IV make the introduction of qualifications necessary. The fact that changing the spatial frequency of the test stimulus so that it is either one octave higher or lower than the adapting stimulus clearly has no effect on the magnitude of the after-effect (Fig. 3). A similar effect has been noted by Campbell and Maffei (1971). They observe that a grating of one frequency can induce a tilt after-effect on a grating of a different frequency or on a single line. This observation is, from the point of view of the model, not in agreement with the fact that particular orientation sensitive units in the cat and monkey striate cortex are responsive to only a limited range of spatial frequencies (Campbell et al., 1969a, b). If the tilt after-effect could be explained in terms of adaptation occurring in orientation selective units in the striate cortex then changing the spatial frequency of the test grating so that there is a difference of one octave between it and the adapting stimulus should result in a decrease of the after-effect to at least half its value in order for the data to show agreement with the behaviour of neurones in the cat and monkey (Campbell et al., 1969a, b) and with human psychophysical data (Blakemore and Campbell, 1969). A second objection to the model was also raised by Campbell and Maffei (1971). They report that the tilt after-effect shows complete interocular transfer. Since the majority of neurones recorded in the striate cortex of the cat and monkey show eye dominance (Hubel and Wiesel, 1959, 1968) it follows that complete interocular transfer of an adaptation effect in these neurones would not be expected.

It is evident from the foregoing discussion that in its present form the hypothesis proposed by Sutherland (1961) will not account for the data available concerning the tilt after-effect. If the model is expanded however a satisfactory explanation of this after-effect could be produced. A class of neurones which responded to orientation independently of the spatial dimensions of the stimulus and which did

not show eye dominance would appear adequately to account for the data. While it might be argued that the invention of a class of neurones for which there is as yet no physiological evidence is unwarranted the current advances in neurophysiology make the testing of the hypothesis feasible. The advantage of retaining a hypothesis which preserves the notion of feature detectors is that such a proposal allows lines and edges to be displaced as a whole thus avoiding some of the difficulties raised by some of the older theories (Sutherland, 1961).

References

BLAKEMORE, C. and CAMPBELL, F. W. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology*, **203**, 237–60.

CAMPBELL, F. W., COOPER, G. F. and ENROTH-CUGELL, C. (1969a). The spatial selectivity of the visual cells of the cat. *Journal of Physiology*, **203**, 223–35.

CAMPBELL, F. W., COOPER, G. F., ROBSON, J. G. and SACHS, M. B. (1969b). The spatial selectivity of visual cells of the cat and the squirrel monkey. *Journal of Physiology*, **204**, 120–21P.

CAMPBELL, F. W. and GREEN, D. G. (1965). Optical and retinal factors affecting visual resolution. *Journal of Physiology*, **181**, 576–93.

CAMPBELL, F. W. and MAFFEI, L. (1971). The tilt after-effect: a fresh look. *Vision Research*, **11**, 833–44.

CAMPBELL, F. W. and ROBSON, J. G. (1968). Applications of Fourier analysis to the visibility of gratings. *Journal of Physiology*, **197**, 551–66.

COLTHEART, M. (1971). Visual feature-analysers and after-effects of tilt and curvature. *Psychological Review*, **78**, 114–21.

DAY, R. H. (1962). Excitatory and inhibitory processes as the basis of contour shift and negative after-effect. *Psychologia*, **5**, 185–93.

GIBSON, J. J. and RADNER, M. (1937). Adaptation, after-effect and contrast in the perception of tilted lines: I. Quantitative studies. *Journal of Experimental Psychology*, **20**, 453–67.

GRANDAGE, A. (1958). Orthogonal coefficients for unequal intervals. *Biometrics*, **14**, 287–9.

HUBEL, D. H. and WIESEL, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *Journal of Physiology*, **148**, 574–91.

HUBEL, D. H. and WIESEL, T. N. (1968). Receptive fields and functional architecture of the monkey striate cortex. *Journal of Physiology*, **195**, 215–43.

KOHLER, W. and WALLACH, H. (1944). Figural after-effects; an investigation of visual processes. *Proceedings of the American Philosophical Society*, **88**, 269–357.

McEWEN, P. (1958). Figural after-effects. *British Journal of Psychology, Monograph Supplement*, No. 31.

OSGOOD, C. E. and HEYER, A. W. (1952). A new interpretation of figural after-effects. **59**, 98–118.

OVER, R. (1971). Comparison of normalisation theory and neural enhancement of negative after-effect. *Psychological Bulletin*, **75**, 225–43.

SCHADE, O. H. (1956). Optical and photoelectric analog of the eye. *Journal of the Optical Society of America*, **46**, 721–39.

SUTHERLAND, N. S. (1961). Figural after-effects and apparent size. *Quarterly Journal of Experimental Psychology*, **13**, 222–8.

ADAPTATION TO PRISMS: DO PROPRIOCEPTIVE CHANGES MEDIATE ADAPTED BEHAVIOUR WITH BALLISTIC ARM MOVEMENTS?

JOHN S. BAILY

Laboratory of Experimental Psychology, University of Sussex

In Experiment I subjects pointed repeatedly at a target viewed through laterally displacing prisms and received terminal visual feedback. In one task the pointing movements were slow (proprioceptively controlled), and in the other they were fast (pre-programmed). In both tasks adaptation proceeded at the same rate and to the same level of performance. Following fast pointing with prisms a large amount of arm-body adaptation was found with slow and fast test movements, while following slow pointing with prisms a large amount of arm-body adaptation was found with slow test movements, but only a small amount with fast test movements. The result suggests that adapted behaviour with pre-programmed movements is not mediated by a proprioceptive change. In Experiment II pointing movements were passive. No arm-body adaptation was found with fast test movements, and, contrary to expectation, only a small amount with slow test movements.

Introduction

When a person practises pointing with one arm at a target viewed through wedge prisms at least two types of adaptation may take place. The first is a change in spatial co-ordination between the eyes and the head (eye-head adaptation), which is probably due, in part at least, to visual asymmetry inherent in the prism transformed visual field (McLaughlin, Rifkin and Webster, 1966). The second is a change in co-ordination between the trained arm and the body (arm-body adaptation), which results from seeing the limb with laterally displaced vision while performing the pointing task (Harris, 1963). This type of adaptation transfers fully when the subject points at non-visual targets such as sound-sources, the contralateral arm or in the direction straight-ahead of the body (Harris, 1963; Craske and Gregg, 1966), and also transfers fully when the subject points 45° to the side of the direction pointed at during prism exposure (Baily, 1970). A third type of adaptation which may occur in some circumstances involves a change in co-ordination between the head and the body (Hamilton, 1964a). The overall change in spatial co-ordination between the eyes and the trained arm (i.e. the change in ipsilateral eye-hand co-ordination) is equal to the *sum* of the component adaptive changes measured in isolation with straight-ahead tests (McLaughlin and Webster, 1967; Wilkinson, 1968). Transfer of adaptation to the contralateral arm is usually mediated by eye-head (or head-body) adaptation (Wilkinson, 1968), and subtraction of the contralateral after-effect from the ipsilateral after-effect

provides a second method for measuring the amount of arm-body adaptation (McLaughlin and Bower, 1965; Wilkinson, 1968). It is with the mechanism underlying arm-body adaptation that we shall be concerned here.

Harris (1963, 1965) postulated that arm-body adaptation is due to a recalibration of the proprioceptive system of the adapted arm, which results in the subject feeling that the arm is located where it looks as if it is when viewed with laterally displaced vision. In support of the proprioceptive change hypothesis, Craske (1966) found that after exposure with fast lateral pointing movements there was an adaptive change in hand/hand co-ordination when the blindfold subject used the non-exposed arm to locate the position of the passively resting exposed arm. The only positional information from the target arm was provided by proprioception.

Although there can be little doubt that a proprioceptive change occurs in arm-body adaptation, it seems that this can only account for adapted behaviour in those situations in which the arm movements are under current proprioceptive feedback control. This is certainly likely to be the case with slow arm movements performed without visual feedback, when it is safe to assume that the limb is moved to where it feels in the correct position. But it is generally agreed that very fast (ballistic) movements are not under feedback control, but are pre-programmed (Keele and Posner, 1967). The simple form of the proprioceptive change hypothesis does not explain how arm-body adaptation could occur with ballistic pointing movements, for the responses would be too rapid for recalibrated proprioceptive feedback to control the direction of movement. In order to adapt in a ballistic pointing task, the subject must learn to pre-programme a new pattern of motor commands to respond to the target. This process will be termed *response learning*.

The proprioceptive change hypothesis could be extended in two ways to explain how proprioceptive recalibration leads to response learning. First, Craske (1968) suggested that a new motor response is pre-programmed due to the error in locating the initial position of the limb according to recalibrated proprioceptive information. Second, it is possible that visual target information is translated into a proprioceptive goal which specifies how the limb will feel when pointing at the target, and a motor response pre-programmed which will achieve this goal.

Alternatively, response learning may be a process which can occur independently of proprioceptive recalibration. Thus there may be two processes involved in arm-body adaptation, which can be identified as changes on the afferent (proprioceptive recalibration) and efferent (response learning) sides of a sensori-motor control system. Adapted behaviour with the trained arm may be mediated by proprioceptive recalibration when the arm movements are controlled by current proprioceptive feedback, and by response learning when the movements are pre-programmed. In Experiment I a method was devised to test whether arm-body adaptation involves one or two processes.

Experiment I

Two pointing tasks with laterally displaced vision were compared. These tasks were identical in all respects except that in one the pointing responses were fast movements, while in the other they were slow movements controlled by proprioceptive feedback (slow and fast exposure tasks). In both tasks terminal visual

feedback was given. The two exposure tasks were compared in terms of the transfer of their after-effects to tests of eye-hand co-ordination performed with slow and fast movements (slow and fast tests).

According to the proprioceptive change hypothesis, adaptation in both tasks will be mediated by proprioceptive recalibration. Although there might be differences between the two tasks in terms of the rate or level of adaptation, the hypothesis predicts that the amount of proprioceptive recalibration produced by a particular task will transfer equally to slow and fast tests.

The hypothesis that arm-body adaptation involves two independent processes predicts that the difference in motor control of slow and fast movements will be critical. Slow exposure will maximize proprioceptive recalibration, since in order to adapt the subject will be forced to compensate for the mismatch between vision and proprioception. Adaptation following slow exposure should transfer to slow tests but not to fast tests. Fast exposure, on the other hand, will maximize response learning, although on the basis of Craske's (1966) result some proprioceptive recalibration can be expected to occur as well. In this case adaptation should transfer to both fast and slow tests.

Method

Subjects

Twenty experimentally naive university undergraduates, 10 male and 10 female, served as subjects. All were right-handed. They were divided into 2 groups containing equal numbers of both sexes. One group underwent exposure with slow movements, the other with fast movements. Subjects received pre-training for 1 h in making slow and fast arm movements in the apparatus.

Apparatus

The semi-circular apparatus (Fig. 1 (a)) was built around a vertical central axis. The subject sat in a chair and was positioned so that the central axis passed through the centre of his head. Head position was stabilized throughout the experiment by using a bite-plate. To measure eye-hand co-ordination subjects pointed below dim target lights (test-targets) viewed in darkness, with no visual feedback from the limbs. The test-targets were 0·25 in. in diameter and 38 in. from the central axis. They were fastened at eye-level to a semi-circular matte black screen. The test-targets used in Experiment I were located 45° left, straight-ahead, and 45° right of the median axis. When not used for pointing responses the hands rested on two bars at waist height set parallel to the frontal plane on the resting surface.

An automated system was used to record the direction and time course of pointing movements. The marking surface was covered with an array of 120 metal strips radiating from the central axis. Each tapered strip was 1° wide and 12 in. long. Metal contacts were attached to the tip of the index finger of each hand. Leads from the contacts were taped to the arms and connected to terminals on the apparatus. The strips of the marking surface were individually connected up to a panel of 120 24 V lamps, which lit up when the strips were touched.

Movement time was recorded with a Dawe electronic timer. All arm movements began from a start switch attached to the underside of the headboard, about 3 in. below the subject's chin. The switch consisted of a 6 in. bar set parallel to the subject's frontal plane, which was pressed back with the dorsal part of the index finger at the beginning of a response. The timer was triggered when the finger left the bar at the start of a movement and was stopped when the finger contact closed the circuit through one of the strips. The panel of lights and the timer were photographed at the end of each response.

Two 15° displacement wedge prisms were mounted base right on a hingeing panel in a mask which covered the subject's face [see Fig. 1 (b)]. Rubber eye-pieces fitted round the orbit of each eye. The size of the circular visual field through the prisms was 30° in diameter. During testing the panel bearing the prisms was swung up from the face, giving a wide field of view. The head was enclosed in a box frame with screens attached to the sides. The subject was unable to see any part of his body or the apparatus except through the apertures in the mask. A hood could be lowered in front of the mask to block all visual input.
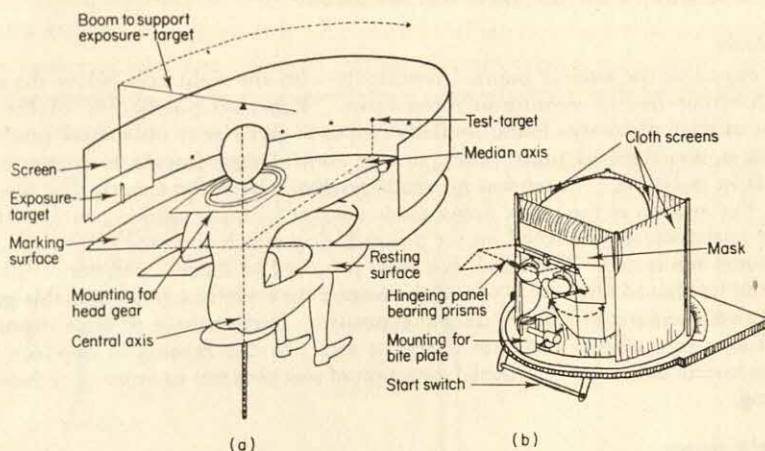

FIGURE 1. The apparatus.

The target for the exposure task (the exposure-target) consisted of a thin vertical 3 in. rod of Perspex with a dim light source at one end. The centre of the rod was at the same level as the test-targets and was positioned 15° right during exposure so as to appear straight-ahead when viewed through the base right prisms. Viewed in the dark the exposure-target appeared as a luminous bar, and with the house-lights on as a white strip. The exposure-target was fixed to the centre of an oblong panel suspended from a boom pivoted above the head, and was placed well to the left of the subject during testing.

The 3 strips of the marking surface 14°, 15°, and 16° right of the median axis were connected to a buzzer which sounded when the subject hit one of these strips during exposure.

Test tasks

Before and after prism exposure eye-hand co-ordination was tested with the left and right arms on the 3 test-targets. In the fast tests the pointing movements were made at the maximal speed the subject could attain. The subject was trained to move the index finger in a straight line from the start switch to the marking surface below the test-target, and not to arch the movement or swing to the side. The movements ended with the arm straight at the elbow. The subject was instructed to look carefully at the test-target and to decide exactly where he was going to point, to attend to the initial position of the finger on the start switch, and to "rehearse" the movement before actually beginning the response. The slow movements were made in two stages. First the subject moved his hand away from the start switch until the arm was straight at the elbow. Then the horizontally extended arm was moved slowly from side to side underneath the test-target, making a series of oscillations of decreasing arc about the position which felt directly below it. The size of the first oscillation was around 30°. This will be termed a *zeroing-in* movement. Having zeroed-in below the test-target the arm was lowered until the index finger contacted the marking surface. The entire movement took at least 3 sec. The subject was trained not to reach out initially directly below the test-target, but always to the side. The object

of the zeroing-in movement was to maximize reliance on current proprioceptive feedback in controlling the response.

The tests were performed in darkness. At the start of each test the subject was presented with a test-target in the dark, and told the arm to use and the type of movement to be made. After each test the arm was returned to its rest position and the house-lights were switched on to allow the experimenter to read from the schedule the details of the next test. At these times the subject was told to keep his eyes closed, and the hood was placed in front of the mask. Performance of the test tasks was not paced.

Exposure tasks

During exposure the subject pointed repeatedly with the right arm below the exposure-target with either fast or zeroing-in movements. Repeated positioning of the arm 15° right of the median plane was found by Baily (1970) to give rise to only small postural after-effects, and it was deemed unnecessary to run control experiments to measure postural after-effects in isolation. There was no visual feedback from the arm during the exposure responses, but as soon as the index finger made contact with the marking surface the house-lights were automatically switched on for 2 sec, during which time the subject kept his arm still and noted his error. The hand was then returned to its rest position. The subject was told that he should attempt to sound the buzzer on every response, and this gave him a clearly defined standard to achieve during exposure. Performance of both exposure tasks was paced at a rate of one response every 15 sec. In the zeroing-in exposure task the subject was instructed to vary his initial direction of reaching out in order to reduce response stereotyping.

Experimental design

Each of the 12 tests of eye-hand co-ordination was given three times before and three times after exposure. Exposure consisted of 45 pointing responses, and was divided into three exposure periods, the first of 25 responses, followed by two periods of 10 responses each. The 12 tests were administered once each after each exposure period. The presentation of the pre- and post-tests within each block of 12 was irregular, with the restrictions that the two arms were used in alternation, and that the same test-target never appeared in two consecutive tests.

Mean changes in eye-hand co-ordination were derived by calculating the difference between the means of the pre- and post-test observations in each test for each subject, and then determining the mean shift in co-ordination in each of the 12 tests in the two experimental conditions.

Results

Performance during exposure

Performance of the two groups during exposure is plotted trial by trial in Figure 2. In both tasks performance was initially in error by the amount expected from the optical displacement. To check that the rate of adaptation was the same in the two tasks a criterion was adopted of five consecutive trials within a range of ±3° of the exposure-target. No significant difference between the two tasks was found (Mann–Whitney test), the mean number of trials to criterion in zeroing-in exposure was 14·7, and in fast exposure 15·3.

After-effects

The contralateral and ipsilateral after-effects are shown in Figure 3. The data for the exposed and non-exposed arms were analysed separately in two three-way analyses of variance in which exposure movement, test movement and test-target position were the main factors.

The analysis of variance for the non-exposed arm showed that none of the main factors nor any of the interactions were significant. This is what would be expected if inter-arm transfer was due to eye-head (or head-body) adaptation, since a systematic change in the apparent locations of visual targets relative to the body should affect all tests of eye-hand co-ordination equally. The mean contra-lateral after-effect was 2·6° (Binomial test $P < 0·001$), and this is taken as the best estimate of the amount of eye-head (or head-body) adaptation that occurred (Wilkinson, 1968).
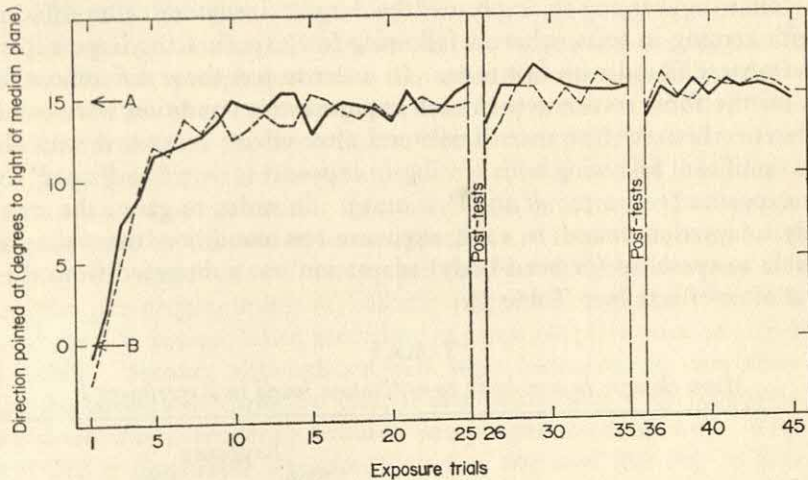

FIGURE 2. Performance during exposure in Experiment I. ——— Zeroing-in exposure; - - - - fast exposure. A—real direction of exposure-target; B—apparent direction of exposure-target.
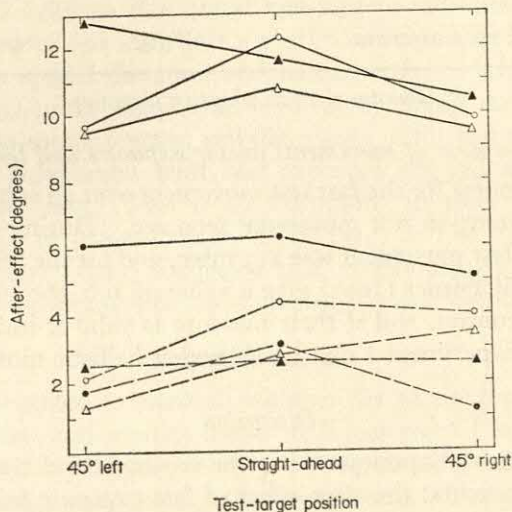

FIGURE 3. Changes in eye-hand co-ordination in Experiment I. ○ Zeroing-in exposure, zeroing-in test; ● zeroing-in exposure, fast test; △ fast exposure, zeroing-in test; ▲ fast exposure; fast test; ——— exposed arm; - - - - non-exposed arm.

The analysis of variance for the exposed arm showed that the type of movement used during testing ($F = 6\cdot29$; df 1,108; $P < 0\cdot05$) and during exposure ($F = 17\cdot78$; df 1,108; $P < 0\cdot001$) were significant factors, as was the interaction between them ($F = 28\cdot36$; df 1,208; $P < 0\cdot001$).   The significant factors were presumably due to the small size of the ipsilateral after-effects measured with the fast test following zeroing-in exposure.   The position of the test-targets was not significant which indicates that the ipsilateral after-effects transferred fully to directions 45° either side of the exposed direction.   This agrees with the results of Baily (1970). The interaction between exposure movement and test movement clearly resulted because following zeroing-in exposure the largest ipsilateral after-effects were found with zeroing-in tests, whereas following fast exposure the largest ipsilateral after-effects were found with fast tests.   In order to test these differences directly the data for the three test-targets in each exposure-test condition were combined. The difference between the mean ipsilateral after-effects measured with the two tests was significant following both zeroing-in exposure ($t = 7\cdot86$; df 29; $P < 0\cdot001$) and fast exposure ($t = 2\cdot77$; df 29; $P < 0\cdot01$).   In order to gauge the amount of arm-body adaptation found in each exposure-test condition the value of 2·6° attributable to eye-head (or head-body) adaptation was subtracted from the mean ipsilateral after-effects (see Table I).

TABLE I

Mean changes in arm-body co-ordination found in Experiment I

| | | Exposure | |
		Slow movement	Fast movement
Test	Slow movement	*8·0°	*7·3°
	Fast movement	*3·3°	*9·0°

* Denotes shift in adaptive direction.

Timecourse of movements during exposure and testing

The mean time-course for the fast test movement over all subjects was 225 msec, and that for the zeroing-in test movement 6·09 sec.   During exposure the mean time-course for the fast movement was 213 msec, and for the zeroing-in movement 6·47 sec.   Keele and Posner (1967) give a value of 190–260 msec for the latency of visual feedback control, and if their measure is valid it indicates that the fast responses made in Experiment I may be classed as ballistic movements.

Discussion

The major finding of Experiment I is the asymmetrical transfer between fast and zeroing-in movements: the after-effect of fast exposure transfers to both fast and zeroing-in movements, while the after-effect of zeroing-in exposure transfers to zeroing-in movements but does not transfer equally to fast movements.   The data indicate that this difference did not result because fast exposure required

quantitatively more adaptation than zeroing-in exposure for performance to become accurate. Nor was the difference due to a larger contribution of eye-head (or head-body) adaptation to the overall change in ipsilateral eye-hand co-ordination following fast exposure.

The finding of asymmetrical transfer supports the hypothesis that two adaptive processes are involved in arm-body adaptation. According to this hypothesis, adaptation in zeroing-in exposure was mediated by proprioceptive recalibration, which showed up as a large shift in arm-body co-ordination in the zeroing-in tests. The finding that this transferred to directions 45° either side of the exposed direction indicates that the recalibration affected a wide range of arm-body orientation. The change in the felt position of the exposed arm did not result in behaviour with fast pre-programmed movements becoming fully adapted as a consequence. This shows that response learning, which must have taken place for behaviour in the fast exposure task to become adapted, does not occur as the result of proprioceptive recalibration, and must be a second adaptive process. There was, however, some transfer from zeroing-in exposure to fast tests, and this was not predicted by the two-process hypothesis. Two reasons for this partial transfer may be suggested. First, proprioceptive recalibration may have *some* effect on the pre-programming of ballistic responses, perhaps when the initial position of the arm is established according to proprioceptive cues, as suggested by Craske (1968). Second, although subjects were instructed to vary their movements in the zeroing-in exposure task, it is possible that with repetition of the response there was a tendency towards pre-programmed control. This might have provided a favourable situation for some response learning to take place. This point is examined further in Experiment II.

Adaptation in the fast exposure task was mediated by response learning, which showed up as a large shift in arm-body co-ordination in the fast tests. The finding that this transferred fully to the lateral test-targets indicates that in response learning the subject does not learn just a particular response, but a more general relationship between spatial direction and motor outflow. Held and Freedman (1963) suggested that prism adaptation involves changing the values of certain constants in the visuo-motor system, and the results of the fast exposure condition support this idea. Adaptation with fast exposure showed a large amount of transfer to zeroing-in tests. This is not surprising in view of Craske's (1966) finding that a change in proprioception takes place in fast movement exposure tasks. Day and Singer (1967) suggested that proprioceptive recalibration may occur as a secondary process which accompanies a motor learning type of process.

Experiment II

Experiment II attempted to establish whether the partial transfer of adaptation from zeroing-in to fast movements found in Experiment I was due to a direct effect of proprioceptive recalibration on pre-programming ballistic responses, or to the occurrence of response learning during exposure. Templeton, Howard and Lowman (1966) showed that an exposure task which required the subject to match disparate visual and proprioceptive inputs was an effective method for producing adaptation with passive arm movements. The two-process hypothesis suggests

that adaptation in this situation is mediated by proprioceptive recalibration, which compensates for the visual-proprioceptive mismatch. It was therefore predicted that if the exposure task used in Experiment I was performed with passive arm movements this would result in as much proprioceptive recalibration as zeroing-in exposure, since both tasks involve making the same inter-sensory judgements. It was expected that the proprioceptive recalibration resulting from passive exposure would lead to a large amount of arm-body adaptation as measured with slow movement tests. With passive exposure it may be assumed that there is no opportunity for any response learning to occur. In this case the amount of transfer of adaptation from passive exposure to fast tests should show the extent to which proprioceptive recalibration has a direct effect on the pre-programming of ballistic responses.

Method

Subjects

Five male university undergraduates served as subjects. All were right-handed.

Apparatus

The apparatus was the same as used in Experiment I except that the test-targets were located 30° left, straight-ahead, and 30° right, and that head position was not stabilized by a bite-plate. During testing the subject was told to keep his head still and aligned with his body, while during exposure head position was stabilized by the subject keeping his face pressed firmly against the eye-pieces of the face mask.
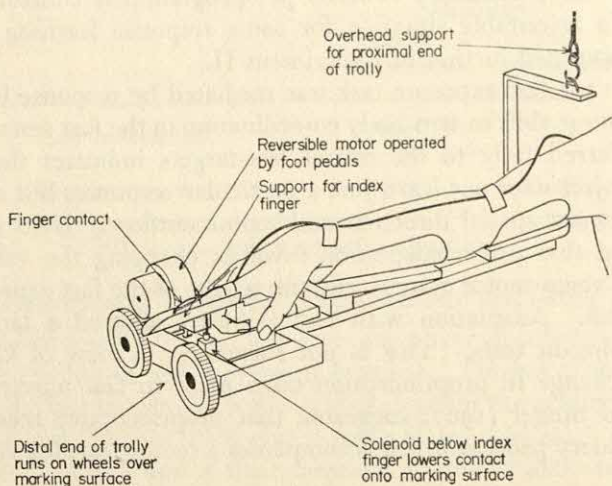

FIGURE 4.   The trolley.

An addition to the apparatus was a trolley for passively moving the subject's right arm. The trolley (see Fig. 4) supported the whole arm in an extended horizontal posture during exposure. The distal end of the trolley ran on wheels over the marking surface and was driven at a speed of 1°/sec by a small reversible motor which was operated by the subject using two foot switches. Disengagement of the motor with a clutch allowed the trolley to be wheeled freely across the marking surface. The right index finger lay outstretched in a trough. A solenoid below the trough lowered a contact onto the marking surface when

the subject operated a switch attached to the resting surface with his left hand, allowing direction of pointing in the trolley to be recorded.

Test tasks

The test tasks were similar to those described for Experiment I, the main difference being a modification to the slow test movement. Instead of reaching out in a straight line in order to get the arm into an extended posture before beginning the zeroing-in part of the movement, the subject now moved the arm at the shoulder so that the hand oscillated from side to side in an increasing arc while the elbow joint was extended. This will be termed a *zero-out/zero-in* movement. The object of zeroing-out was to decrease the possibility of pre-programmed control. The subject was trained to take at least 5 sec to make this movement.

Exposure task

The exposure task was the same as used in Experiment I (prisms base right, exposure-target 15° right), and was performed with passive arm movements. The subject was instructed to keep the muscles of the arm and shoulder as relaxed as possible, and frequent reminders about this were given. Before each exposure response the trolley motor was disengaged and the right arm pushed by the experimenter to a start position between 15° and 30° to the side of the exposure-target. Start positions were alternated left and right. The passive movement strategy was to approach the exposure-target continuously from the start position, to carry on until the arm was felt to have overshot, and then to go back to what felt to be the correct position. Performance was not paced during exposure and movement times were not recorded. The average rate of performance was 2–3 responses per min, which was rather slower than in Experiment I.

It is most probable that the muscles of the arm in the trolley were not completely relaxed, although without EMG records it is not possible to gauge the degree of relaxation. However, the lack of an EMG record is not a serious difficulty. The criterion of passive movement used in previous experiments which have investigated possible differences in adaptation between active and passive arm movements (Held and Hein, 1958; Pick and Hay, 1965; Templeton, Howard and Lowman, 1966; Singer and Day, 1966a,b) was that the limb was moved by an external agent while the muscles were in a relatively relaxed state, not that the limb showed total neuromuscular silence.

Experimental design

The experimental paradigm was the same as used for Experiment I.

Results

The initial error in exposure was 2·2° less than expected from the optical displacement. The mean number of exposure trials to a criterion of five consecutive trials within an error range of ±3° was 16·0, almost the same as in Experiment I. The mean timecourse of the zero-out/zero-in movements was 7·36 sec, and of the fast movements 180 msec.

Contralateral and ipsilateral after-effects are shown in Figure 5. The data for the two arms were submitted to a three-way analysis of variance in which test-target position, hand used, and test movement were the main factors. The hand used was the only significant factor ($F = 14·68$; df 1,48; $P < 0·001$), which indicates that the means of the ipsilateral and contralateral after-effects were significally different. The only significant interaction was between the hand used and test movement ($F = 6·62$; df 1,48; $P < 0·05$), which clearly resulted because the largest contralateral after-effects were found with fast tests while the largest ipsilateral after-effects were found with zero-out/zero-in tests.

The contralateral after-effects measured with zero-out/zero-in and fast tests were not significantly different (*t*-test). The mean contralateral after-effect was 0·7°, which was significantly different from zero ($t = 2·23$; $df\,29$; $P < 0·05$), and may be taken as an index of the amount of eye-head (or head-body) adaptation that occurred.
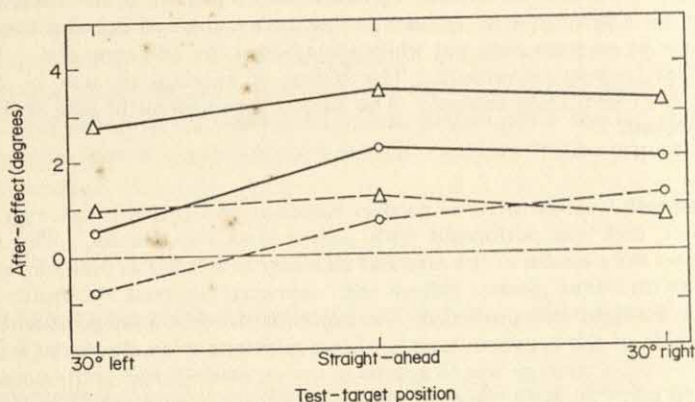

FIGURE 5. Changes in eye-hand co-ordination in Experiment II. △ zero-out/zero-in test; ○ fast test; —— exposed arm; - - - - non-exposed arm.

The mean ipsilateral after-effect measured in the fast tests was 1·6°, and this was not significantly larger than the mean contralateral after-effect of 0·7° (*t*-test). The mean ipsilateral after-effect measured in the zero-out/zero-in test was 3·3°, and this was significantly larger than the mean contralateral after-effect ($t = 7·59$; $df\,14$; $P < 0·001$). Subtracting the mean contralateral after-effect gives a value of 2·6° for the amount of arm-body adaptation measured with zero-out/zero-in tests. The finding that this transferred 30° either side of the exposed direction shows that a generalized change in arm-body co-ordination had occurred.

Discussion

The most striking result of Experiment II is that the amount of arm-body adaptation measured with the zero-out/zero-in tests was very much less than predicted from the assumption that passive and zeroing-in exposure tasks would give rise to the same amount of proprioceptive recalibration. In view of this it is not valid to conclude that the lack of transfer to the fast tests shows that proprioceptive recalibration has no effect on the pre-programming of ballistic movements. Further experiments with the trolley indicate that the small change in arm-body co-ordination found with the zero-out/zero-in tests was not the result of a postural after-effect (Baily, in preparation), and it may be concluded that this shift was due to proprioceptive recalibration. The question arises why the amount found was so small? Two reasons may be suggested.

First, the assumption underlying Experiment II may have been wrong: proprioceptive calibration may not occur unless the arm is moved actively, even when the exposure task involves visual-proprioceptive matching. Hamilton

(1964b) suggested that activity in the motor system may serve as a trigger or switch that in some manner enables correlations to be made between visual and proprioceptive estimations of positions of members of the body. The importance of active movement as a critical determinant of adaptation has been emphasized by Held and Freedman (1963). In this connection it is worth pointing out that previous experiments which claimed to have demonstrated adaptation with passive arm movement (Templeton, Howard and Lowman, 1966; Singer and Day, 1966a,b) did not include tests which would have shown whether arm-body adaptation was involved. Adaptation in these experiments could have been due to eye-head adaptation, which can result simply from viewing an asymmetrical visual field (McLaughlin, Bifkin and Webster, 1966). In the present experiment the difference between ipsilated and contralateral after-effects shows that a small amount of arm-body adaptation was produced by passive exposure. However, this could have resulted because the arm muscles were not completely relaxed during exposure (Pick and Hay, 1965).

A second possibility is that a large amount of proprioceptive recalibration does occur with passive exposure, but is conditioned to the trolley. The trolley may provide strong tactual cues such that the change in the felt position of the arm is intimately bound up with how the arm feels while in the trolley. Although it is not usual for adaptation to be conditioned to cues associated with the exposure task—if it were, after-effects would always be very small—instances of this are on record (Kohler, 1964).

References

BAILY, J. S. (1970). Prism adaptation and motor control. Unpublished doctoral thesis, University of Sussex.

CRASKE, B. (1966). Change in transfer function of joint receptor output. *Nature*, **210**, 764–5.

CRASKE, B. (1968). An analysis of human adaptation to prismatically displaced vision. Unpublished doctoral thesis, Durham University.

CRASKE, B. and GREGG, S. J. (1966). Prism after-effects: identical results for visual targets and unexposed limb. *Nature*, **212**, 104–5.

DAY, R. H. and SINGER, G. (1967). Sensory adaptation and behavioural compensation with spatially transformed vision and hearing. *Psychological Bulletin*, **67**, 307–22.

HAMILTON, C. R. (1964a). Intermanual transfer of adaptation to prisms. *American Journal of Psychology*, **77**, 457–62.

HAMILTON, C. R. (1964b). Studies on adaptation to deflection of the visual field in split-brain monkeys and man. Unpublished doctoral thesis, C.I.T.

HARRIS, C. S. (1963). Adaptation to displaced vision: visual, motor or proprioceptive change? *Science*, **140**, 812–3.

HARRIS, C. S. (1965). Perceptual adaptation to inverted, reversed and displaced vision. *Psychological Review*, **72**, 419–44.

HELD, R. and FREEDMAN, S. J. (1963). Plasticity in human sensorimotor control. *Science*, **142**, 455–62.

HELD, R. and HEIN, A. V. (1958). Adaptation of disarranged eye-hand coordination contingent upon re-afferent stimulation. *Perceptual and Motor Skills*, **8**, 87–90.

KEELE, S. W. and POSNER, M. I. (1967). Processing of visual feedback in rapid movements. *Journal of Experimental Psychology*, **77**, 155–8.

KOHLER, I. (1964). The formation and transformation of the perceptual world. (Trans. H. Fiss.) *Psychological Issues*, **3**.

McLAUGHLIN, S. C. and BOWER, J. L. (1965). Selective intermanual transfer of adaptive effects during adaptation to prism. *Psychonomic Science*, **3**, 69–70.

McLAUGHLIN, S. C., RIFKIN, K. I. and WEBSTER, R. G. (1966). Oculomotor adaptation to wedge prisms with no part of the body seen. *Perception and Psychophysics*, **1**, 452–8.

McLAUGHLIN, S. C. and WEBSTER, R. G. (1967). Changes in straight-ahead eye position during adaptation to wedge prisms. *Perception and Psychophysics*, **2**, 34–44.

PICK, H. L. and HAY, J. C. (1965). A passive test of the Held reafference hypothesis. *Perceptual and Motor Skills*, **20**, 1070–2.

SINGER, G. and DAY, R. H. (1966a). The effects of spatial judgements on the perceptual after-effect resulting from prismatically transformed vision. *Australian Journal of Psychology*, **18**, 63–70.

SINGER, G. and DAY, R. H. (1966b). Spatial adaptation and after-effect with optically transformed vision. Effects of active and passive responding and the relationship between test and exposure responses. *Journal of Experimental Psychology*, **71**, 725–31.

TEMPLETON, W. B., HOWARD, I. P. and LOWMAN, A. E. (1966). Passively generated adaptation to prismatic distortion. *Perceptual and Motor Skills*, **22**, 140–2.

WILKINSON, A. (1968). Adaptation to displaced sensory input. Unpublished doctoral thesis, Oxford University.
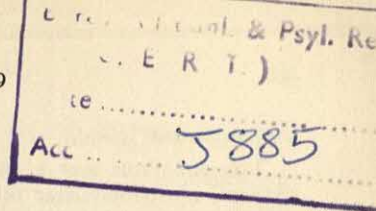
RETROACTIVE CONTOUR ENHANCEMENT: A NEW VISUAL STORAGE EFFECT

L. G. STANDING† AND P. C. DODWELL

Department of Psychology, Queen's University, Kingston, Ontario

A test stimulus (a visual form) which is below recognition threshold when flashed briefly against a steady background field can be raised to complete discriminability if the background field is terminated and replaced by darkness within about 100 msec of the test flash. There must therefore be an efficient storage process for the apparently "invisible" form. The phenomenon is shown to occur under a variety of conditions. It appears not to be a simple visual masking phenomenon. The tasks used are forced-choice, and the phenomenon provides a new demonstration of visual storage effects which previously have generally been reported only for supraliminal visual test stimuli.

Introduction

Many recent studies show that the perceived duration of a brief visual stimulus may greatly exceed its physical duration (Sperling, 1960; Averbach and Coriell, 1961; Eriksen and Collins, 1967; Keele and Chase, 1967; Haber and Standing, 1969). Often a flash of 50 msec or less persists visually for over 200 msec; the visual storage time may be increased by dark-adapting the subject. These studies, which use various techniques to measure the storage effect (partial report, metacontrast, visual summation, and direct ratings) all deal with stimuli presented above the recognition threshold.

The present study demonstrates a visual storage effect for "subliminal" stimuli, by means of an apparently novel visual phenomenon which we term retroactive contour enhancement. This phenomenon is produced by presenting a test contour which is below the recognition threshold when it is flashed briefly against a steady background field. When the background is terminated shortly after the test stimulus offset, the test contour becomes fully visible; there is retroactive enhancement of its sensory trace.

In Experiment I the retroactive enhancement effect is established, using a group of subjects; in Experiment II some of the properties of the effect are explored in more detail with a single observer.
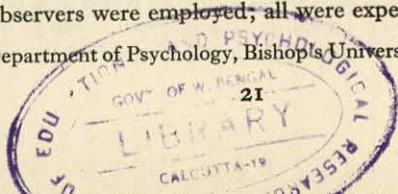
Experiment I

Method

Subjects

Ten emmetropic observers were employed; all were experimentally naive.

† Present address: Department of Psychology, Bishop's University, Lennoxville, Quebec, Canada.

Apparatus and stimuli

The apparatus was a two-channel Maxwellian-view optical system. Two Sylvania R1131C glow-modulator tubes, each controlled by an American Electronic Laboratories stimulator (model 104A), provided light sources with rise and decay times of less than 1 msec. The two square stimulus fields were concentric, each subtending 6° 11′, and there was a small central fixation point.

The luminance of one stimulus field was set at either 22 or 2·2 mL; this provided a steady masking background field. A sub-threshold test stimulus was provided by a 20 msec exposure of the second field; this field always contained one of two high-contrast black test patterns (the letter S or L), which both subtended 46′ vertically. The letter was a transparency placed in the parallel portion of the second optical channel. The luminance of the test field was either 9 or 1·4 mL for the corresponding high and low mask luminance conditions. Testing was performed in a moderately lighted room, with a background luminance of 0·2 mL in the apparatus at all times.

Procedure

Appropriate levels of luminance of the test field (probe) were first determined so that the probability of correctly discriminating the two test contours was approximately 0·6, when the probe was flashed against a steady mask field (chance performance having a 0·5 probability). Two determinations were made, with the mask at 22 mL and at 2·2 mL, yielding the two probe luminances already mentioned.
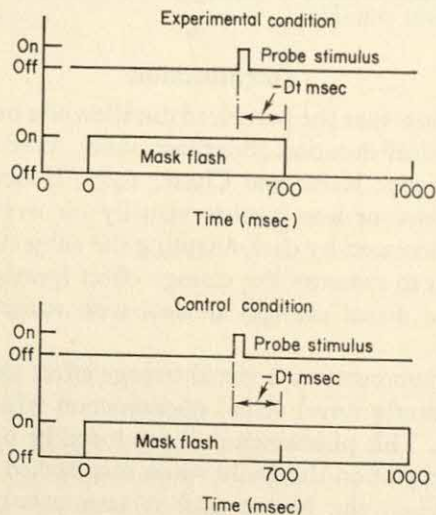

FIGURE 1. Stimulus sequence for experimental and corresponding control conditions.

The subjects were then required to discriminate between the two test contours under experimental or control conditions. Under experimental conditions, the probe was flashed −Dt msec before, or +Dt msec after the termination of the mask field (which lasted 700 msec). Under the control condition, the probe was flashed at the same delay after the onset of the mask field as in the corresponding experimental condition, but the mask duration was prolonged to 1000 msec; the probe was thus presented essentially on a steady background of the mask field, since the mask offset occurred much later. The experimental and control paradigms are shown in Fig. 1.

Subjects were tested in a 2 × 2 × 8 factorial design (experimental vs. control condition, high vs. low mask intensity, and 8 Dt values ranging from −300 to +100 msec.) Each was given a block of 10 trials under each treatment combination, the treatments being selected in

random order. The probe luminance was always set at the appropriate value (as deter-
mined by the pretest) for the mask luminance being employed.

On each trial one of the two possible test contours was presented, in random sequence, and
the subject was required to make a forced-choice discrimination between them. No cues
from slide-changing were available, since slides were changed on every trial, using duplicate
sets of test stimuli. Subjects were informed in advance of the two possible stimuli, and told
that one would be presented just before or after the offset of the background field on each
trial. Each subject was tested in a single session lasting about 2 h.

Results

The mean probability of correct discrimination between the two test contours
is shown in Figure 2. The increase in correct responses which is evident under
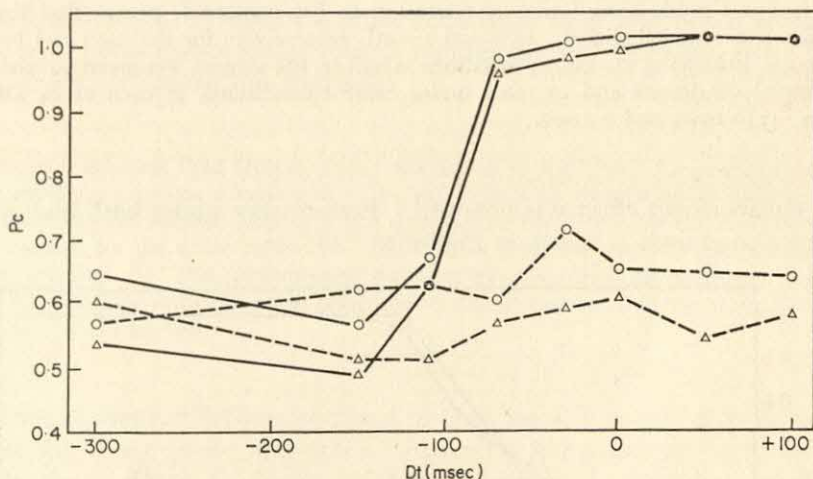experimental conditions is highly reliable ($F = 194$; $df = 1,279$; $P < 0.01$).

FIGURE 2. Probability of correct response (Pc) in contour identification task, Experiment I
group data. The mask offset occurs at $Dt = 0$ under experimental conditions: under control
conditions the mask terminates 300 msec after $Dt = 0$.
    High luminance: ○ —— ○ Experimental; ○ - - ○ Control.
    Low luminance: △ —— △ Experimental; △ - - △ Control.

A significant effect of mask luminance was found ($F = 9.99$; $df = 1,279$;
$P < 0.01$), while the interaction between mask luminance and testing conditions
(experimental vs. control) was not reliable ($F = 2.0$; $df = 1,279$; $P < 0.05$).
This indicates that the probe was slightly more discriminable under the high-
luminance condition, for both experimental and control readings, no doubt
because the threshold value obtained in the pretest was slightly high. We may
conclude that the change in detectability in the experimental condition is essentially
the same under high and low mask-luminance conditions.

Reliable differences in performance between subjects were also found ($F =
4.58$; $df = 9,279$; $P < 0.01$).

Experiment II

To explore the enhancement effect found in Experiment I in more detail,
without pooling data from several subjects, the procedures of Experiment I were

repeated, using a single well-practised observer, with several different test and mask stimuli. The same apparatus was used throughout.

(a) *Replication of Experiment I with lower probe luminance*

Aim

This experiment was intended to verify the results of Experiment I with a more precise measurement of the effect and with the test stimulus definitely below threshold luminance for recognition (i.e., below the luminance at which the probability of correct discrimination under control conditions was 0·5 for the single observer).

Method

The test and mask stimuli were the same as in Experiment I, except that the probe luminance was reduced slightly, to 8 and 1·2 mL respectively for the high and low mask luminances. Following the same procedures as before, the subject was given 40 trials under experimental conditions and 20 trials under control conditions, at each of 14 Dt values between −110 msec and 0 msec.

Results

The enhancement effect was obtained. Performance under both high and low luminance conditions is shown in Figure 3.
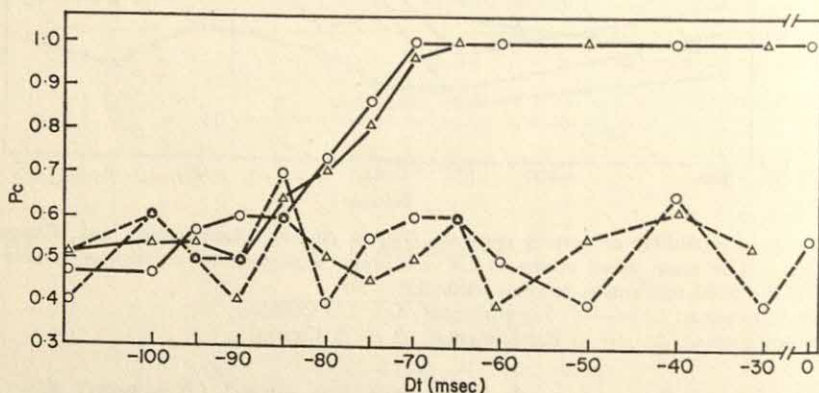

FIGURE 3. Performance (Pc) in Experiment II(a), for a single subject.
High luminance: ○ —— ○ Experimental; ○ - - ○ Control.
Low luminance: △ —— △ Experimental; △ - - △ Control.

(b) *Variations in mask luminance with probe luminance constant*

Aim

The effects of varying the mask luminance above and below the "low" value used in Experiment I were studied, with the probe luminance constant at the subthreshold value used in Experiment II(a) (1·2 mL).

Method

The mask luminance was successively set at −0·2, +0·4 and +0·7 log units relative to 2·2 mL; for each value, the observer's performance was measured at Dt values between −30 and −150 msec as in Experiment II(a).

Results

Increases in mask luminance caused corresponding decreases in the —Dt values at which the complete retroactive enhancement effect was obtained; in other respects the enhancement functions were very similar to those shown in Figure 3. At the lowest mask luminance level, enhancement was complete for — Dt values up to 70 msec, at the intermediate level up to 60 msec, and at the highest level of mask luminance enhancement was complete only for —Dt values of 40 msec or less.

(c) *Effects of visual noise, and of decreased probe luminance*

Aim

This experiment measures the enhancement effect when the mask field is a random brightness field, with the probe luminance at or just below recognition (50% correct performance under control conditions), and when the probe luminance is 1·0 log unit below this threshold.

Method

A random brightness field (Julesz, 1964) was added to the "low" luminance mask field (thereby reducing its average luminance to 1·4 mL from 2·2), and the probe luminance was raised to 4·8 mL, the highest level at which detection of the probe did not rise above 50% when presented on the noise field under the control condition (mask offset extended by 300 msec, see Fig. 1). The enhancement function was then obtained as before. A second function was obtained under identical conditions to the first except that the probe luminance was reduced by 1·0 log unit.

Results

The enhancement effect was obtained in both cases, but over a much smaller range for the lower probe luminance. Again the enhancement functions were similar to those shown in Figure 3; at the higher probe luminance the effect ceased to occur at —Dt values greater than 90 msec, for the lower probe luminance at values above 40 msec.

(d) *Effects of visual noise in the masking field*

Aim

Following the surprisingly strong enhancement effect obtained in Experiment II(c), despite the presence of visual noise in the mask field, it was decided to replicate study II(a) (low mask-luminance conditions) with visual noise added to the mask field but without any other change.

Method

The same noise field used in Experiment II(c) was added to the low-luminance mask field, reducing the average value to 1·4 mL. Probe luminance was unchanged at 1·2 mL. Performance was measured as before.

Results

The enhancement effect was obtained over approximately the same interval when the visual noise was present as when the mask field was blank but at a somewhat higher luminance.

(e) *Relationship between off-effect masking and enhancement*

Aim

The relation between the contour enhancement effect and the off-effect masking produced by offset of a background field found in previous studies (Crawford, 1947; Dodwell and Standing, 1970) was studied within the same experimental situation.

Method

Two separate functions were obtained. The first was for the detection of a 20 msec, 1° square probe flash; this was at 0·9 mL, the luminance yielding 90% detectability under control conditions. The second function was for the detection of a contour superimposed on the above probe flash; the probe flash was now presented on every trial (instead of half the trials, as for the first function), but the contour was present on only half the trials. The subject was required to report on each trial whether the contour (the L used in Experiment I) was present or not.
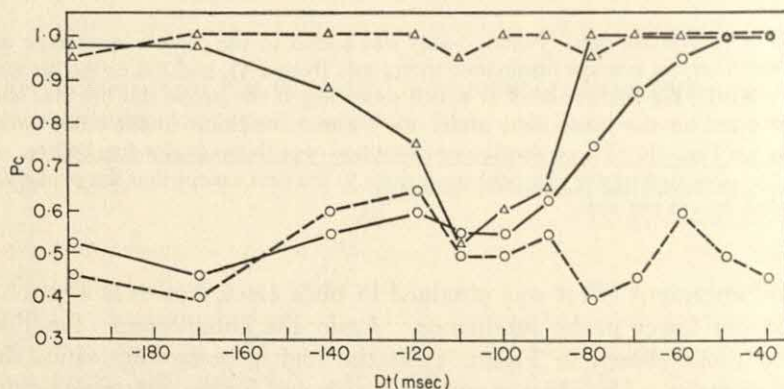

FIGURE 4. Performance (Pc) in flash detection (masking) and contour detection (retroactive enhancement) measured within one experiment. For the masking function, each data point represents 50 experimental or 25 control readings; for the enhancement function the data points each represent 40 readings.

Enhancement: ○ —— ○ Experimental; ○ - - ○ Control.
Masking: △ —— △ Experimental; △ - - △ Control.

Results

The data are shown in Figure 4. One function shows the usual off-effect masking phenomenon, as found by Crawford and others. The second function shows that the detectability of a test contour superimposed on the probe flash gives the characteristic retroactive enhancement effect obtained previously for discrimination between the two possible contours.

Discussion

The results of Experiment I demonstrate that the recognition of a contoured test stimulus may be retroactively raised, by terminating a background adapting field within approximately 100 msec after the presentation of the test stimulus. Per-

formance under these conditions is compared with a control condition where the test contour is flashed against a steady background of the adapting field, and is not recognized.

The present phenomenon could be described as a failure of the adapting field to "mask" the test stimulus unless it is prolonged more than 100 msec beyond the test stimulus offset. This is not incorrect, but fails to capture the full significance of the finding. "Masking" is itself an ambiguous term, since it can refer to the masking of a target by a steadily-presented masking field, as in our control conditions, or to the interference between rapidly-presented sequences of visual events. To call our effect a failure to mask, then, would be correct in the first sense of masking. Possibly off-effect interference-type masking plays some role in preventing access to the stored contour information but it obviously is not the major factor; target recognition (under enhancement conditions) and detection (under masking condition) are quite dissimilar even at small values of $-Dt$, as is shown in Figure 4. This point is reinforced by the fact that introducing visual noise into the mask has a relatively weak effect on the enhancement function. Smith and Schiller (1966) have shown that introduction of visual noise into a mask flash normally increases the amount of interference-type masking considerably. Systematic changes in the contour characteristics of the mask were not explored; this would certainly be a worthwhile exercise. The Julesz field used was of finer grain than the contours of the probe, and no doubt more efficient masking (in the first sense) could be obtained with different types of mask, for instance a noise field of coarser grain or one containing other sorts of contour information.

A major difference between our situation and the normal interference type masking paradigm is this: in a masking paradigm one is normally interested in the fact that two temporally and/or spatially discrete events interfere with each other, in a sense apparently may overlap each other (cf. Dodwell and Standing, 1970, for example), whereas in our situation one event (the mask) completely overlaps another (the probe) both spatially and temporally, and yet information in the second event is still preserved intact. We do not claim that subjects could not detect the *presence* of the probe in the enhancement condition, only that they could not recognize the spatial pattern in it unless the mask was terminated within a short period. While it is true that access to the spatial pattern is prevented by the steadily-presented mask, it would seem odd to claim that an interfering-type event (offset of the mask) should itself make the pattern information available. In this sense we suggest that our finding is genuinely different from the usual type of masking phenomenon.

It seems reasonable to call it a storage phenomenon, and apparently it has not been reported before; we suggest that "retroactive contour enhancement" is an appropriate name for it. It is related to the general idea of "short-term visual storage" as conventionally defined (e.g. Keele and Chase, 1967) since in both cases there is quite long perceptual persistence following a brief stimulus flash. It differs from previous demonstrations of visual storage in that those demonstrations generally have used stimuli which were above recognition threshold at the time of presentation. In one case (Standing, Haber, Cataldo and Sales, 1969), a subliminal storage effect was inferred; the stimulus trace (below threshold for individual

flashes) was "activated" by flashing it repetitively, rather than by the onset of darkness as in the present case. Another phenomenon which bears some relation to the present effect has been reported by Sperling (1960b). Under specific conditions of relative luminance, a test flash which is followed immediately by a mask flash is invisible, but a (negative) after-image of the test flash is clearly discernible. As with the present study, a rather weak stimulus trace must be persisting (over roughly the same time interval), to be "reactivated" by a change in background luminance.

Two points deserve special emphasis. The first is that the contoured probe stimulus is below recognition threshold (and can be well below, as shown in Experiment II(c)) and, secondly, the effect is demonstrated in a forced-choice task for both experimental and control conditions. Thus the presence of sensory storage processes may be inferred somewhat more confidently than in most studies of supraliminal short-term visual storage, since to infer the storage process in the latter case generally involves an additional assumption about processing. For example, in the partial-report experiment of Sperling (1960a), it is necessary to assume that the subject is reading out the marked stimuli from a true visual image rather than from some more abstract representation of the test stimulus in memory. While such criticism may not be crucial, there is a potential flaw in the evidence for storage when stimuli are presented supraliminally, and it is satisfactory to confirm the existence of storage processes by another, converging, operation. It is particularly satisfying to find (Experiments II(a-e)) that the storage effect is very robust under a variety of experimental conditions, and does not depend on a special choice of stimulus parameters. It may be noted in passing that there have been other reports of behavioural effects of "subliminal" patterns in a masking situation (e.g. Fehrer and Raab, 1962) but these did not involve subsequent phenomenal recovery of the pattern. Our effect is probably much more closely related to the delay of contour formation studied by Kahneman (1965), although the two can hardly be identical, in view of our Experiment II(d), which involves presentation of contour on contour. This same experiment makes it unlikely that a simple luminance-summation principle can explain the effect.

The studies reported as Experiment II show that the most important determinant of storage is the respective luminances of the probe and mask, the former increasing the duration of storage, the latter decreasing it. They also show that the phenomenon is different from off-effect backward masking as reported by Crawford (1947), and by Dodwell and Standing (1970). The latter phenomenon (as shown again here in Experiment II(e)) is represented by a U-shaped detectability function, rather than by the monotonic function found for the storage effect. There is little doubt that a test flash presented at threshold just before the offset of a background field is masked, whereas when a test flash containing a visual form is presented with the form at or just below the recognition threshold, the storage effect is obtained.

We believe that this phenomenon presents interesting new possibilities for studying the temporal properties of the visual system.

References

AVERBACH, E. and CORIELL, A. S. (1961). Short-term memory in vision. *Bell System Technical Journal*, **40**, 309–28.

CRAWFORD, B. H. (1947). Visual adaptation in relation to brief conditioning stimuli. *Proceedings of the Royal Society, London*, B134, 238–302.

DODWELL, P. C. and STANDING, L. G. (1970). Studies of visual backward masking and a model for the Crawford effect. *Acta Psychologica*, **32**, 31–47.

ERIKSEN, C. W. and COLLINS, J. F. (1967). Some temporal characteristics of visual pattern recognition. *Journal of Experimental Psychology*, **74**, 476–84.

FEHRER, E. and RAAB, D. (1962). Reaction time to stimuli masked by meta-contrast. *Journal of Experimental Psychology*, **63**, 143–7.

HABER, R. N. and STANDING, J. G. (1969). Direct measures of short-term visual storage. *Quarterly Journal of Experimental Psychology*, **21**, 43–54.

JULESZ, B. (1964). Binocular depth perception without familiarity cues. *Science*, **145**, 356–62.

KAHNEMAN, D. (1965). Exposure duration and effective figure-ground contrast. *Quarterly Journal of Experimental Psychology*, **17**, 308–14.

KEELE, S. W. and CHASE, W. G. (1967). Short-term visual storage. *Perception and Psychophysics*, **2**, 383–6.

SMITH, M. C. and SCHILLER, P. H. (1966). Forward and backward masking: a comparison. *Canadian Journal of Psychology*, **20**, 191–7.

SPERLING, G. (1960a). The information available in brief visual presentations. *Psychological Monographs*, **74** (Whole No. 498).

SPERLING, G. (1960b). Negative afterimage without prior positive image. *Science*, **131**, 1613–4.

STANDING, L. G., HABER, R. N., CATALDO, N. and SALES, B. D. (1969). Two types of short-term visual storage. *Perception and Psychophysics*, **5**, 193–6.

NEUROPSYCHOLOGICAL EVIDENCE OF VISUAL STORAGE IN SHORT-TERM MEMORY TASKS

ELIZABETH K. WARRINGTON AND T. SHALLICE†

National Hospital, Queen Square, London, W.C.1

Auditory and visual presentation of verbal material were compared in a single patient having an auditory verbal S.T.M. deficit. A Peterson short-term forgetting experiment and an immediate memory span task are reported. Striking differences in performance related to modality of input were obtained. Auditory short-term forgetting was more rapid, whereas with visual presentation short-term decay functions were relatively normal. With visual presentation there was no evidence of acoustic confusion errors but there was some evidence of visual confusion errors. The findings are interpreted in terms of a separate post-perceptual visual S.T.M. system.

Introduction

A specific impairment of verbal short-term memory was described in a patient (K.F.) by Warrington and Shallice (1969). There was a marked reduction of span for auditorily presented verbal items (digits, letters and words), whereas both the span for similar material presented visually and the visual span of apprehension were less impaired. It was demonstrated that the defective auditory immediate memory span could not be accounted for by faulty auditory perception nor in terms of a motor speech defect. It was argued that there was a modality specific impairment of auditory verbal S.T.M. storage; and that recall of visually presented verbal material was, at least in part, from a separate post-perceptual visual S.T.M. store.

K.F.'s performance on certain two-component memory tasks was examined to provide more "positive" evidence for a short-term memory defect (Shallice and Warrington, 1970). Using the Peterson procedure (auditory presentation) the normal rapid decay function was not observed and this finding was interpreted in terms of an absence or near-absence of the S.T.M. component of the task. One shortcoming of this experiment was his relatively poor performance with the zero delay condition. The three-letter word stimuli were beyond K.F.'s. immediate memory span. It is of interest to consider whether there would be rapid forgetting of verbal stimuli within his span. Adequate recall with no delay and abnormally rapid forgetting under conditions of delayed recall would provide strong support for the view that K.F.'s deficit is one of S.T.M. storage.

The Peterson procedure also provides a convenient paradigm for comparing

† Present address: Department of Psychology, University College, Gower Street, London, W.C.1.

short-term forgetting of auditorially and visually presented test stimuli. In normal subjects very similar but not identical decay functions are obtained on a variety of S.T.M. tasks for both modalities of presentation (Murdock, 1966; Margrain, 1967; Conrad and Hull, 1968; Craik, 1969). It is therefore of interest to extend the original findings of relatively spared visual memory span, to a visual short-term memory forgetting task. Different decay functions depending on modality of input would provide further evidence for separate post-perceptual stores for auditory and visual material.

The evidence for acoustic coding processes operating with visual presentation (Conrad, 1964) has led to models of information processing involving transfer of the visual input into a post-perceptual auditory store (Sperling, 1963; Neisser, 1967). An iconic or sensory store indicating the availability of information in a visual stimulus in the post-exposure period for up to 1 or 2 sec is generally accepted (Sperling, 1967; Broadbent, 1970). However, if there were a separate post-perceptual visual store, one would predict the absence of or lack of reliance on acoustic coding processes under those conditions when recall is direct from such a visual store. It is thus relevant to consider the frequency of acoustic errors made by K.F. in relation to modality of input when the duration of the information in "store" exceeds that possible with iconic storage.

The aim of the present study was to obtain further evidence for (1) a specific defect of auditory verbal S.T.M. store, and (2) the existence of a post-perceptual S.T.M. visual store. In this investigation the patient K.F. was tested on the Peterson procedure using as stimuli single letters and trigrams of letters with both auditory and visual presentation. In addition, an error analysis of an auditory and visual immediate memory span task is presented.

Experiment I—The Peterson Procedure

Short-term forgetting of verbal items with auditory and visual presentation was examined using the Peterson technique.

Method

Test stimuli comprised single letters, pairs of letters, and trigrams. All letters of the alphabet, excluding I and O, were used equally often in a random sequence, the only constraint being no repetition within a pair or trigram. The test stimuli were hand-written in black ink in upper case on a 2 × 2½ in. white card for manual visual presentation. A 1 item/sec presentation rate was used for both visual and auditory presentation; in the latter the letters were spoken by the experimenter. In the delay conditions, an intervening task (reading aloud singly presented random numbers as fast as possible) was employed. Each trial block was preceded by three discarded trials at a constant interval of 15 sec.

Three conditions of this experiment were done independently, using different delay periods.

Condition (1)

Recall of single letter stimuli was tested after delays of 0, 5, 10 and 15 sec. K.F. was tested in blocks of 20 trials, five for each delay, in random order. Two blocks of visually presented items and two blocks of auditorially presented items were completed, making a total of 80 trials. Visual and auditory presentation of trial blocks were alternated.

Condition (2)

Recall of pairs of letter stimuli was tested after delays of 0, 5, 10, 15 and 30 sec. K.F. was tested in blocks of 10 trials, two at each delay in random order. Twelve blocks, six with auditory presentation and six with visual presentation, in alternate order, were completed.

Condition (3)

Recall of letter trigrams was tested after delays of 0, 5, 10, 15, 30 and 60 sec. K.F. was tested in blocks of 12 trials, two at each delay, in random order; 10 trial blocks, alternating visual and auditory presentation, were completed.

Results

The percentage of single letters recalled (disregarding order), for both auditory and visual presentation, is shown in Figure 1, for each delay interval. Under these delay conditions there was no forgetting with visual presentation. However, with auditory presentation though there was perfect recall with zero delay, there was a
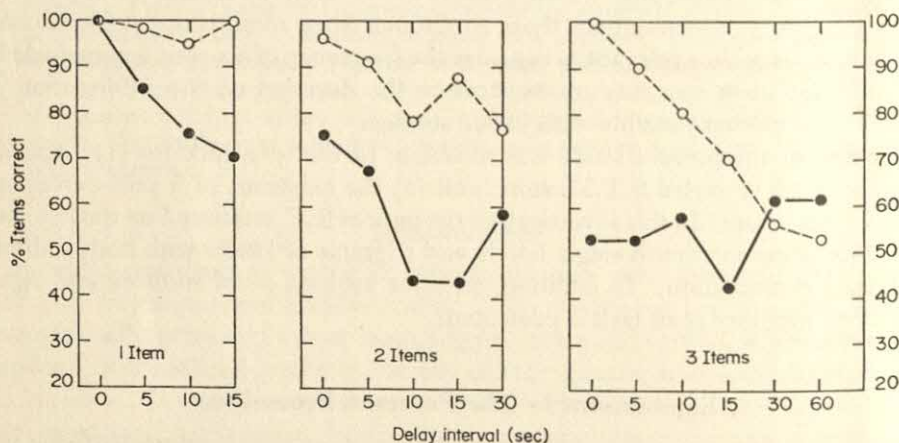

FIGURE 1. Peterson—Pre-operative results. ○ - - - Vision. ● —— audition.

decrement in performance related to the length of the interval. The percentage of items recalled with two-letter and three-letter stimuli is shown in Figure 1, for both auditory and visual presentation. In the two-item condition there was a greater decrement with increasing delays for the auditory stimuli than the visual stimuli. The percentage recalled with auditory stimuli at a zero delay interval is relatively good, but clearly shows a mild impairment. The decay curve for visually presented trigrams appears relatively normal, in marked contrast to the auditory curve where K.F.'s performance with zero delay is so bad that virtually no further forgetting was recorded. These findings will be discussed in conjunction with those of Experiment II.

Experiment II—Post-operative Peterson Procedure

Shortly after the above results were obtained, K.F. underwent surgery for the treatment of his epilepsy. A craniotomy was performed, and the anatomical extent of scar tissue and electrically abnormal tissue established, but no radical intervention

was possible (Warrington, Logue and Pratt, 1972). K.F. was retested one month post-operatively, at which time his speech and immediate memory span were somewhat improved. He was now able to repeat two items fairly reliably but had considerable difficulty with three items. It therefore seemed of interest to repeat the Peterson procedure, both because of the possibility of obtaining better performance with the zero delay conditions, and of establishing whether or not there is any relationship between immediate memory span and short-term forgetting.

Method

The testing procedure was the same as described in the first experiment. Recall of single letters, two letters and three letters was tested both with visual and auditory presentation after delays of 0, 5, 10, 30 and 60 sec. K.F. was tested in blocks of 10 trials, two for each delay interval. There were eight blocks of trials for each number of test stimuli and for each mode of presentation, making a total of 48 trial blocks. The three-letter recall conditions were done on a different occasion from the single and double letter recall conditions; the latter were alternated. Trial blocks of visual and auditory presentation were alternated throughout.
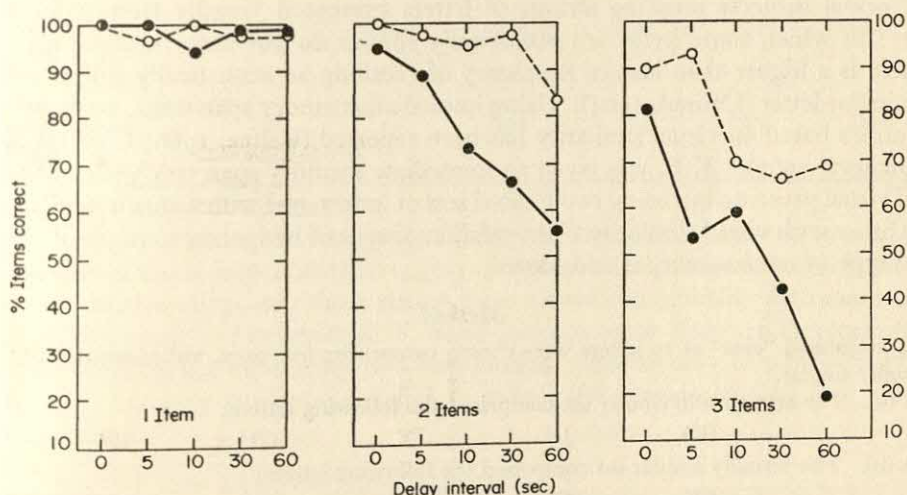

FIGURE 2.   Peterson—Post-operative results.   ○ - - - Vision.   ● —— audition.

Results

The percentage of letters recalled with visual and auditory presentation are shown separately for one, two and three letter stimuli in Figure 2. Recall of single letters shows no decrement in either modality under these delay conditions. Recall of two letters presented visually shows very little decrement with increasing delay, whereas recall of two letters presented auditorially is nearly accurate at zero delay, but clearly shows a decrement with increasing intervals. Again with three-letter strings, the forgetting curve with visual presentation is relatively normal, whereas with auditory presentation performance with zero delay is poor and shows further decrement with delay. There is thus a consistent difference according to the modality of input, a greater decrement in recall being found with auditory stimuli than visual stimuli.

Taking the results of both these experiments, there are two critical results, one-letter stimuli pre-operatively and two-letter stimuli post-operatively. In both these conditions, in spite of perfect or near-perfect registration of the stimuli as indicated by the zero delay condition, forgetting is much more rapid for auditory stimuli than for the same material presented visually. Second, it can be seen that the asymptotic level (if such exists) is not a constant but is related to the modality of input. In only one condition, three-letter stimuli pre-operatively, were the asymptotes for auditory and visual presentation similar. In addition, if one calculates the "capacity" of the system at the longest delay condition (60 sec) in terms of number of letters per trial recalled, one obtains estimates post-operatively of 0·96, 1·12 and 0·57 for one, two and three letters presented auditorially and of 0·96, 1·69 to 2·06 for one, two and three letters presented visually. The 0·96 value may result from a ceiling effect, but nevertheless the number of items per trial recalled at 60-sec delay varies greatly between the two modalities.

Experiment III—Immediate Memory Span—Error Analysis

Normal subjects recalling strings of letters presented visually from a limited "set" in which some items are acoustically similar do not make random errors. There is a higher than chance frequency of recalling an acoustically similar than dissimilar letter (Conrad, 1964). Using immediate memory span tasks, no evidence of errors based on visual similarity has been reported (Calfee, 1966; Cimbalo and Laughery, 1967). K.F. was given an immediate memory span task with auditory and visual presentation using two limited sets of letters, one with acoustic similarity, the other with visual similarity. The relation of type of item error to mode of input and type of confusability is considered.

Method

Two limited "sets" of 10 letters were chosen comprising five pairs, either acoustically or visually similar.

Set (a). The acoustically similar set comprised the following letters:

BD   JK   IY   QU   SX

Set (b). The visually similar set comprised the following letters:

EF   OQ   KR   IL   MZ

As far as possible, visual confusions were minimized in set (a) and acoustic confusions in set (b).

K.F. was tested post-operatively on an immediate memory span task under four conditions, all of which were done on different occasions. For all four conditions, a 2 per sec presentation rate was used. K.F. was instructed to write his responses in correct serial position on ruled paper, the set of 10 letters in alphabetical order being available.

(1) Auditory presentation of acoustically similar letters. Strings of three letters from set (a) were presented in blocks of 40 trials, five blocks in all, making a total of 200 trials. The letters were chosen randomly, and used equally often, the only constraint being no repetition within a string.

(2) Auditory presentation of visually similar letters. Strings of three letters of set (b) were again presented in blocks of 40 trials, with a total of 200 trials. No letter was repeated in the same string.

(3) Visual presentation of acoustically similar letters. Strings of four letters from set (a), each written in black ink on a white card 2 × 2½ in., were presented in blocks of 40 trials, with a total of 200 trials.

(4) Visual presentation of visually similar letters. Strings of four letters from set (b), as above, a total of 200 trials presented in blocks of 40 were completed.

Results

Under the conditions of this experiment, in addition to errors of item recall, K. F. made a great number of order errors and therefore only item information was included in the analysis so as to facilitate a precise comparison between the modalities. Each item error in the response was classified as either non-random if the other item of the pair had appeared in the string, or random (see Table I). By

TABLE I

Distribution of type of errors

| | Visual confusion strings | | | Acoustic confusion strings | | |
	Random	"Visual"	Total	Random	"Acoustic"	Total
Auditory Presentation (String of 3)	167	44	211	130	71	201
Visual Presentation (String of 4)	168	55	227	169	47	216

using three-letter strings in the case of auditory presentation and four-letter strings in the case of visual presentation a roughly comparable error rate has been obtained. For statistical analysis only those strings were considered in which both items of a pair did not occur and in which K. F. made only one error. One can then calculate precisely the chance expectancy of a non-random response as 1 in 7 for strings of three (i.e. auditory presentation) and as 1 in 6 for strings of four (i.e. visual presentation).

TABLE II

Observed and expected frequency of errors

| | | Visual confusion strings | | | Acoustic confusion strings | | |
		Random	"Visual"	Total	Random	"Acoustic"	Total
Auditory	O	48	6	54	59	25	84
Presentation	E	46	8		72	12	
Visual	O	65	16	81	57	9	66
Presentation	E	68	13		55	11	

The total number of errors (after the above elimination procedure) in each category for each condition of the experiment is given in Table II, together with the number of each type of error expected by chance. A chi-squared test was used to compare the observed values with the expected values. There was a significant

increase in acoustic confusions with auditory presentation ($\chi^2 = 17 \cdot 0$; $P < 0 \cdot 001$; $df = 1$), in marked contrast to visual presentation of the same letters, where the frequency of acoustic confusions is very close to chance expectation ($\chi^2 = 0 \cdot 4$). On the other hand, there was no significant increase of visual confusion errors with visual presentation ($\chi^2 = 0 \cdot 8$).

Experiment IV—Visual Immediate Memory Span—Error Analysis

Qualitative observations have suggested that K.F. uses a visual "code", yet no evidence to support this notion was obtained in the above experiment. In the previous experiment K.F. frequently had to "guess" the fourth letter at random. If his introspections are valid then a small visual confusion effect could well have been masked. In order to overcome this possibility a procedure was devised by which a "guessing" response would not contaminate genuine non-random responses. This was done by embedding a limited set of letters in the whole alphabet in such a way that estimates of chance probability could be made from a comparison of errors within the limited set.

Method

A set of four pairs of visually similar but acoustically dissimilar letters was chosen:

OQ   MW   PR   YX

Strings of three letters comprising two from the above set and one chosen at random from the remaining letters of the alphabet were presented visually and K.F. was instructed to write the letters in correct serial position on ruled paper. 100 trials were completed at a 2 per sec rate and 100 trials at a 1 per sec rate. In order to allow an accurate calculation of the chance probability, only responses with one item error were considered in the analysis. Errors for stimuli other than the critical set of eight letters were discarded, together with responses which were not letters in the critical set. The remainder, that is (1) a visual confusion error or (2) a non-visual confusion error from the critical set, were considered in terms of frequency in relation to chance expectation. By chance a visual error should occur on 1 in 6 of the analysed trials.

TABLE III

Distribution of visual errors

	Random errors Observed	Random errors Expected	"Visual" errors Observed	"Visual" errors Expected	Total Observed
(1) ½ sec presentation rate	19	27	14	6	33
(2) 1 sec presentation rate	26	31	11	6	37
(3) combined ½ & 1 sec presentation rate	45	58	25	12	70

Results

The frequency of these two types of error together with the chance expectancy, for each rate of presentation, is given in Table III. There is a significantly high incidence of visual errors with the $\frac{1}{2}$ sec and 1 sec presentation rates ($x^2 = 15.7$; $P < 0.001$; $df = 1$, and $x^2 = 4.9$; $P < 0.05$; $df = 1$, respectively): combining both sets of data $x^2 = 18.1$ ($P < 0.001$; $df = 1$).

Discussion

The data reported in this study clearly support the view that K.F. has a specific defect of auditory verbal short-term memory, and that there is a separate post-perceptual visual store. His short-term forgetting of auditory material is much more rapid than for visual material. Moreover, his immediate memory span with visual presentation, unlike that of normal subjects and his own acoustic presentation performance, is not subject to acoustic confusions.

In our original experiment (Shallice and Warrington, 1970) using the Peterson procedure there was a significant trend for longer intervals to produce lower recall scores, which was no longer significant if the zero condition was omitted. This slight trend was interpreted in terms of a much reduced auditory S.T.M. capacity together with a normal L.T.M. component. This interpretation is open to the objection that there was inadequate recall at zero delay. The present data provide much more satisfactory evidence for the reduced S.T.M. capacity hypothesis. There are two critical conditions, one-item stimuli pre-operatively and two-item stimuli post-operatively. K.F.'s performance is close to 100% accuracy with zero delay, indicating adequate registration, yet forgetting is much more rapid than in normal subjects tested under similar conditions (see e.g. Murdock, 1961; Melton, 1963).

Visual and auditory input with the Peterson procedure give markedly different forgetting curves. For one, two and three items, there is a greater decrement for auditory presentation than for visual presentation. It is argued that this difference reflects recall directly from a visual store. The possibility that it is due to easier access to an auditorially encoded S.T.M. for visual than for auditory material is not tenable, since with visual presentation acoustic confusion errors did not exceed chance expectation. Given that these results do reflect forgetting from a visual store, in view of the decline lasting over some 30 sec, it is plausible to assume that the memory trace is post-perceptual in nature.

The concept of a pre-perceptual or iconic trace, which extends beyond the presence of the stimuli by some msec, is a powerful explanation of certain phenomena of visual recall (Sperling, 1960). Posner's (1969) evidence of a visual store with a duration of 1-2 sec may not in fact reflect the same process, as masking stimuli are effective in Sperling's normal subjects but not in Posner's. However, recent experiments in normal subjects comparing recall with auditory and visual presentation suggest effects with durations up to 10 sec (Margrain, 1967; Murdock and Walker, 1969) which would be difficult to account for on a pre-perceptual iconic store hypothesis. A recent study of short-term memory while shadowing (Kroll, Parks, Parkinson, Bieber and Johnson, 1970) goes still further and provides

evidence of a post-perceptual visual store with a duration in the region of 20–30 sec. This latter finding is in close agreement with the present study which suggests the possibility of a visual store with comparable duration, if not capacity, to the auditory store.

Performance on an auditory immediate memory task is superior to a visual immediate memory task in normal subjects. The opposite is the case for K.F. It is therefore likely that in normal subjects the function of the visual S.T.M. system is masked by the superior capacity of the auditory S.T.M. system, whereas in our patient K.F., with a much reduced auditory S.T.M., the visual S.T.M. system is operating in relative independence. One possibility is that these results are of limited generality in that K.F. has compensated for his auditory defect by developing abnormally good visual skills, but this is difficult to assess. The observation that a further patient with selective impairment of auditory verbal S.T.M. had an auditory span of 3–4 and a visual span of 8 immediately post-operatively, where it can be assumed that insufficient time for compensation or relearning had elapsed (Warrington, Logue and Pratt, 1972), makes this view implausible.

The Peterson procedure has been widely regarded as a two-component task—one showing rapid forgetting over the first few seconds, the other more stable, represented by the asymptote. On this view one would expect the long-term component (i.e. the asymptotic part of the curve) to be much the same for both modalities. However, the level of asymptote with visual and auditory presentation is comparable in only one of the six conditions, namely three items pre-operatively. In addition, an analysis in terms of "capacity" at the 60 sec delay condition gives widely varying estimates which would not be predicted by the stable L.T.M. component hypothesis. The situation remains unclear, since not only in our original pre-operative data (Shallice and Warrington, 1970) using three three-letter word stimuli but also with three-letter stimuli pre-operatively there was less decrement with increasing delay intervals than post-operatively, when his overall performance on span tasks and short-term forgetting tasks was superior. It is tentatively suggested that pre-operatively his S.T.M. functions were so poor that he attempted to code the input in terms appropriate for L.T.M. and so showed no decline with longer intervals. Thus this pre- and post-operative difference with three-item stimuli presented auditorially may reflect a change in strategy. The result obtained by Baddeley and Warrington (1970) in amnesic subjects, whose decay functions were normal at all points of the curve, including 30 sec and 60 sec, is further evidence that the level of an asymptote may not be determined just by L.T.M. Hence the relative contributions of L.T.M. and S.T.M. at the longer intervals in the Peterson task remain obscure.

It is argued that if K.F. is recalling information from an independent visual short-term store, on this view it must follow that with visual presentation of acoustically similar lists, the frequency of acoustic confusion errors would not be greater than chance. This was indeed found; with visual presentation the incidence of acoustic errors did not exceed chance, using a limited set of letter stimuli which resulted in a significant increase in acoustic errors with auditory presentation.

In the hope of elucidating the "visual code", K.F. was tested with visually confusable letters. In the first error analysis, no evidence of visual errors with

visual presentation was obtained. However, in a second procedure designed to eliminate random response and highlight visual confusions, a significantly greater than chance frequency of visual errors was recorded. Visual coding processes have been shown to operate in situations where recall or recognition are tested within 2 sec of the offset of the stimulus (Posner, 1968). The present findings offer tentative evidence that post-perceptual visual coding is also in terms of shape.

Finally, it is worth considering the present findings in relation to information processing models. Most current theorists favour a model whereby visual information after being perceived and identified has to be transferred to an auditory store for immediate storage, from which it is recalled (Sperling, 1963; Conrad, 1964; Neisser, 1967). The evidence of acoustic errors in immediate memory tasks with visual presentation is vital to this position. However, the results of the present study suggest that this is not necessarily the case. Independent recall from a visual store is indicated not only by K.F.'s superior performance with visual presentation, but also by the absence of acoustic errors on the visual immediate memory span task. The present data are consistent with Sperling's (1967) concept of a visual buffer store, but would suggest that it is visual rather than a motor movements programme, of relatively long duration compared with that of the icon, and with direct access for retrieval purposes.

References

BADDELEY, A. D. and WARRINGTON, E. K. (1970). Amnesia and the distinction between long and short-term memory. *Journal of Verbal Learning and Verbal Behaviour*, **9**, 176–89.

BROADBENT, D. E. (1970). Psychological aspects of short-term and long-term memory. *Proceedings of the Royal Society, London B*, **175**, 333–50.

CALFEE, R. C. (1966). How learning is affected by change in subject matters: sources of interference in verbal learning. *Technical Report*, University of Wisconsin.

CIMBALO, R. S. and LAUGHERY, K. R. (1967). Short-term memory: effects of auditory and visual similarity. *Psychonomic Science*, **8**, 57–8.

CONRAD, R. (1964). Acoustic confusion in immediate memory. *British Journal of Psychology*, **55**, 75–84.

CONRAD, R. and HULL, C. L. (1968). Information, acoustic confusion and memory span. *British Journal of Psychology*, **12**, 1–6.

CRAIK, F. I. M. (1970). Modality effects in short-term storage. *Journal of Verbal Learning and Verbal Behaviour*, **8**, 658–64.

KROLL, N. E. A., PARKS, T., PARKINSON, S. R., BIEBER, S. L. and JOHNSON, A. L. (1970). Short-term memory while shadowing. *Journal of Experimental Psychology*, **85**, 220–4.

MARGRAIN, S. A. (1967). Short-term memory as a function of input modality. *Quarterly Journal of Experimental Psychology*, **19**, 109–14.

MELTON, A. W. (1963). Implications of short-term memory for a general theory of memory. *Journal of Verbal Learning and Verbal Behaviour*, **2**, 1–21.

MURDOCK, B. B. (1961). The retention of individual items. *Journal of Experimental Psychology*, **62**, 618–25.

MURDOCK, B. B. (1966). Visual and auditory stores in short-term memory. *Quarterly Journal of Experimental Psychology*, **18**, 206–11.

MURDOCK, B. B. and WALKER, K. D. (1969). Modality effects in free recall. *Journal of Verbal Learning and Verbal Behaviour*, **8**, 665–76.

NEISSER, U. (1967). *Cognitive Psychology*. New York: Appleton-Century-Crofts.

POSNER, M. I. (1970). Abstraction and the process of recognition. In BOWER, G. H. and SPENCE, J. T. (Eds), *The Psychology of Learning and Motivation : Advances in Research and Theory*. New York, London: Academic Press.

SHALLICE, T. and WARRINGTON, E. K. (1970). The independent functioning of verbal memory stores: A neuropsychological study. *Quarterly Journal of Experimental Psychology*, **22**, 261–73.

SPERLING, G. (1960). The information available in brief visual presentations. *Psychological Monograph*, **74**, no. 11.

SPERLING, G. (1963). A model for visual memory tasks. *Human Factors*, **5**, 19–31.

SPERLING, G. (1967). Successive approximations to a model for short-term memory. *Acta Psychologica*, **27**, 285–92.

WARRINGTON, E. K., LOGUE, V. and PRATT, R. T. C. (1972). The anatomical localisation of selective impairment of auditory verbal short-term memory. *Neuropsychologia*. In press.

WARRINGTON, E. K. and SHALLICE, T. (1969). The selective impairment of auditory verbal short-term memory. *Brain*, **92**, 885–96.

SERIAL POSITION EFFECTS IN THE RECALL OF PICTURE SEQUENCES

GILLIAN COHEN

Institute of Experimental Psychology, University of Oxford

Sequences of 10 pictures of familiar objects were presented for immediate recall, and typically bow-shaped serial position curves were obtained. The effects of redundant stimulus suffixes, and of redundant stimulus prefixes, which could be either extra pictures, or written words (the names of similar familiar objects) were examined. In the prefix conditions, extra words produced a general decrement over positions 1–4, as compared with a control condition. Picture prefixes had no effect. In the suffix conditions, redundant words affected positions 9 and 10, whereas the main effect of redundant pictures was at positions 8 and 9. The differential serial position effects are taken to reflect the presence of two memory codes, a verbal one and a visual one, which vary in trace-strength or accessibility at different list positions.

Introduction

In verbal learning tasks, the study of serial position curves, with and without irrelevant additional items, has been illuminating, and has yielded valuable insights into the coding and storage mechanisms which are employed. The present experiment utilized the same technique in order to examine the characteristics of memory for sequences of pictures.

Numerous experiments on the serial learning of auditorily presented items such as letters, digits and nonsense syllables have studied the effects of redundant suffix and prefix items on retention at different serial positions. This method has been successful in demonstrating that the primacy effect (the initial part of the serial learning curve), and the recency effect (the terminal part of the curve), are differentially sensitive to such interference. These results have been interpreted by Crowder and Morton (1969) as reflecting two types of store. According to their theory, the recency effect stems from a pre-recognition sensory register type of store, and retention of the terminal items in this store is markedly disrupted by stimulus suffix items. The primacy effect results from those items which have been recognized and transferred to STM, and it may be slightly reduced by stimulus prefix items (Dallett, 1965), but is unaffected by suffixes. Crowder and Morton have failed to obtain recency effects with visual presentation of alphanumeric material; they conclude that the visual sensory register is too short-lived to mediate any such effects. Recency effects have been found however, in some other experiments with visually presented pictures (Potter and Levy, 1969; Paivio, Rogers and Smythe, 1968). Pictures may remain available longer than letters or words before being verbally re-coded, or lost.

One object of the present study, then, was to determine whether the serial

position curves typically found with auditory material, could be obtained with visual presentation when the stimuli were pictures rather than letters, digits or words.

The experiment tested retention of sequences of pictures of familiar objects, and employed redundant stimulus suffix and stimulus prefix items which could be either verbal or pictorial. In general it has been found that interference is greatest when redundant items are physically similar to the list items, that is, when they are presented in the same modality and the same voice. Cross-modal suffix effects were absent in an experiment by Morton and Holloway (1970), and the prefix effects were eliminated by a change of voice in an experiment by Neisser, Hoenig and Goldstein (1969). Although Parlett (1968) found some evidence that redundant items interfered more when they were of the same class as list items, Morton (1968) obtained strong suffix effects with redundant digits in a letter list. Crowder (1970) concludes that the nature of the suffix item is unimportant; the physical channel over which it arrives is crucial. If this is so, the redundant words and the redundant pictures in the present experiment should exert an equal effect, since all the items are visually presented. If the nature of the redundant items is critical, extra pictures should cause more interference than extra words since they are more similar to the list items.

Alternatively, it is possible that redundant items cause more interference when they coincide with the way in which the list items are coded, stored or tagged for retrieval. This hypothesis gives rise to several different predictions:

(1) If the list pictures are retained as images, then the redundant pictures would cause more interference.

(2) If the pictures are named and coded verbally, the extra words would cause more interference.

(3) If the pictures are retained, as Paivio et al. (1968) have suggested, in a dual coding, both verbal and visual, a more complex result would be predicted. The two codes might well vary in relative availability or in trace-strength. In this case, extra pictures should impair retention at list positions where retrieval is primarily via the visual code, and extra words should cause more disruption at list positions where items are retrieved via the verbal code. Then, given that a serial position effect does occur in memory for picture sequences, and that it is differentially affected by the two kinds of redundant item, it should be possible to infer the nature of the memory code operative at different stages of the sequence.

Method

Materials

The pictures were black and white line drawings of familiar objects (e.g. chair, cow, ball, clock, feather) made up into 24 × 36 mm transparencies. There were 100 different pictures which constituted the list items, and an additional 20 different pictures which appeared only as redundant items. 20 words which were the names of different familiar objects were printed with Letraset and similarly mounted as slides. The words were only used as redundant items.

Procedure

Each subject worked through six conditions, and was presented with 10 lists in each condition. The lists consisted of 10 items, so that each of the 100 pictures appeared once only in each condition. The pictures were randomly re-ordered in each condition. They were

presented by means of a Kodak projector with carousel, and back-projected onto a screen to yield a size of 9×6 in. at a viewing distance of 4 ft. Presentation was controlled at a 2 sec rate throughout by a unit of Behavioural and Digital Research Logic. The six conditions were as follows:—

PWB (Pictures with Words Before). The first two slides were of words which constituted redundant prefixes. One blank slide intervened before the list of 10 pictures. Recall began immediately after the final picture terminated.

PPB (Pictures with Pictures Before). The first two slides were of redundant pictures. The to-be-learned list of 10 pictures followed an intervening blank as in the previous condition.

PWA (Pictures with Words After). The 10 picture sequence occurred first, and was followed by one intervening blank slide and then two words which constituted redundant suffixes. Recall followed immediately after the last of the redundant words.

PPA (Pictures with Pictures After). This condition was exactly as the *PWA* condition, except that the redundant suffixes were additional pictures.

PBB (Pictures with Blanks Before). This was a control condition in which three blank slides preceded the picture sequence. This condition is therefore equated with *PWB* and *PBB* in respect of any temporal uncertainty which might arise as a result of the delayed onset of the critical items.

PBA (Pictures with Blanks After). This was a further control condition in which three blank slides intervened between the end of the list and the beginning of recall. Comparison of the *PWA* and *PPA* conditions with this control should then yield differences attributable only to the interference, since delay in recall is equated.

The 10 item lists are longer than those usually employed in serial learning experiments, and two redundant items are added, although the normal practice is to add only a single one. These departures from the method of previous studies were necessary because pictorial material is more easily learned, and the differences between conditions would otherwise have been obscured by ceiling effects.

18 subjects were tested, and the order of conditions was balanced across subjects. The subjects were university and pre-university students of both sexes, and were paid for their services. Each attended for two sessions and worked through three conditions in separate blocks at each session. Sessions lasted about an hour.

Each condition was explained to the subject at its outset. He was told that he must look at the redundant items, but that he would not have to remember them. Recall was written. Subjects were allowed to recall the list in any order, but had to arrange items in the correct order of presentation by writing them in numbered slots on the recall sheet. This method was preferred to a strictly serial recall so as to avoid obscuring recency effects by output interference.

Results

A response was scored as correct only if it was in the correct position. The total numbers of errors made by all subjects in the six conditions are graphed in Figures 1 and 2.

A three-way analysis of variance, list position × type of interference (picture blank or word) × position of interference (before or after), was performed on the error scores. The results yielded a significant effect of list position, $F = 46.13$; $df = 9,18$; $P < 0.001$; of type of interference, $F = 3.87$; $df = 2,18$; $P < 0.05$; and of before/after, $F = 5.05$; $df = 1,18$; $P < 0.05$. The interaction of list position with type of interference was also just significant, $F = 2.34$; $df = 18,18$; $P < 0.025$.

Differences between the conditions were examined further by means of a Scheffé test. The results, significant at $P < 0.025$, were as follows:—

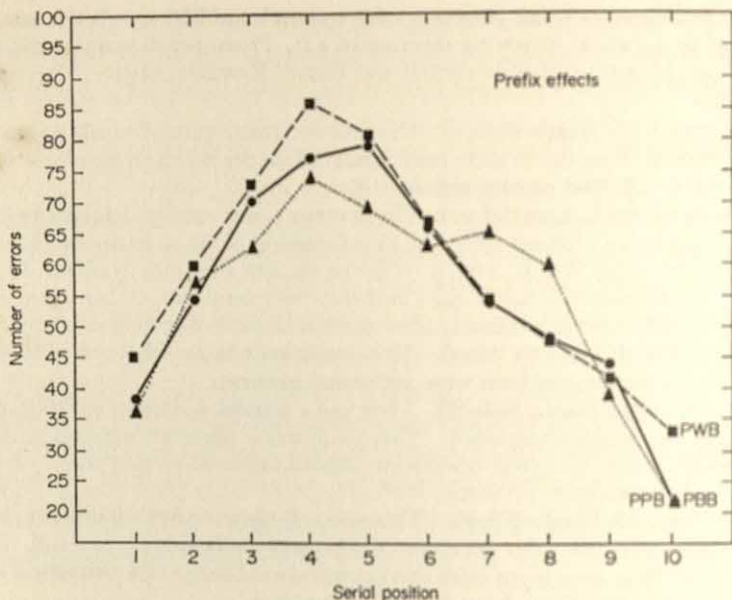$$PBB = PPB < PWB = PBA = PPA < PWA$$

FIGURE 1. The total number of errors for all subjects at each serial position (out of 180 responses), in conditions *PWB* (redundant word prefixes), *PPB* (redundant picture prefixes) and the control condition *PBB* (delayed onset, but no interference).
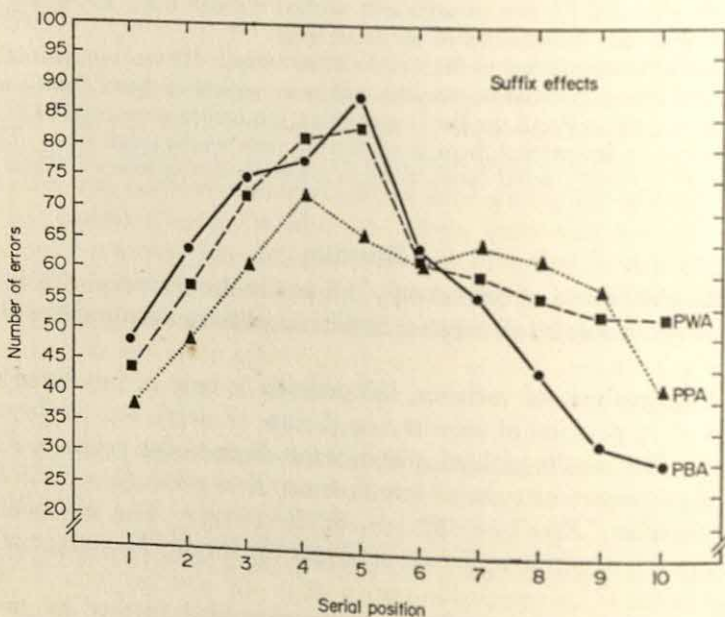

FIGURE 2. The total number of errors for all subjects at each serial position (out of 180 responses), in conditions *PWA* (redundant word suffixes), *PPA* (redundant picture suffixes) and the control condition *PBA* (delayed recall, but no interference).

Differences between serial positions were then tested by the Wilcoxon matched-pairs signed-ranks test. Each of the interference conditions was tested against its control. The results can be summarized as follows, the level of significance being $P < 0.01$.

(1) *Prefix effects.* There was no significant effect of either word or picture prefixes at any particular list position, although the *PWB* condition yielded a higher over-all score; there is a slight general elevation over earlier list positions.

(2) *Suffix effects.* The *PWA* condition produced significantly more errors than *PBA* in list positions 9 and 10, as well as a higher over-all score. The *PPA* condition produced more errors than *PBA* in positions 8 and 9; errors were also increased in positions 7 and 10, as can be seen in Figure 2, though not to a significant extent. This decrement was partly offset by the anomalous improvement at positions 1 to 5, so that the over-all error scores for *PPA* were not higher than for *PBA*.

The levels of significance for these results are not high. However, exactly the same differences were apparent when the results of the first 12 subjects were analysed, and were still present when the full number of 18 subjects had been tested, which indicates that the observed effects are replicable.

Most of the error responses were omissions or transpositions. Of the 11 intrusion responses which occurred, 10 were the names of the redundant pictures, and only one was a redundant word.

Discussion

The ordered recall of picture sequences does exhibit a serial position effect, the curves obtained being very similar to those hitherto confined to the recall of auditory material. However, the present experiment differs from previous studies in a number of important respects. In particular, the material presented for recall is pictorial rather than alphanumeric; and the method of recall permits the subject to reproduce the series by unloading his memory in any order. The production of the typical bow-shaped curve with visually presented lists may well be dependent on either or both of these factors.

The interference effects are also similar to those obtained in the Crowder and Morton studies, in that they are mostly limited to the terminal list positions. Since in this experiment the two types of interference, verbal and pictorial, affect the serial position curve differentially, although both are presented in the same physical channel, the nature of the redundant items must be affecting performance. But the similarity of list items to the redundant items cannot be critical, because the interference effects are not confined to those conditions with pictorial redundant items.

The results obtained are best interpreted in terms of a complex relationship between the memory coding system, and the nature of the redundant items. Again the interaction of serial position with type of interference forces the conclusion that the memory code cannot be uniform over the entire series. Suppose, however, that each picture in the list is retained in both a visual and a verbal code. It is plausible that, as presentation of new items and rehearsal of earlier items proceed, the relative availability or trace-strength of the two codes varies with list position. Redundant items might then cause more interference when they coincide with the

memory code which is stronger, or more available for retrieval at a particular list position. (It is not necessary for this argument to adopt any one of the current models of memory. The only necessary assumption is that visual and verbal codes may co-exist in varying states of accessibility.)

The prefix effects showed that redundant words, but not redundant pictures, caused a slight elevation of the earlier part of the error curve. The inference is therefore that the items in the first part of the sequence are retrieved via the verbal code.

In the suffix conditions, both word and picture suffixes caused decrements in performance over the terminal part of the curve, suggesting that both verbal and visual codes may be utilized for retrieval at this stage. The major effect of the picture suffixes occurs slightly earlier in the list than that of the word suffixes. Recall appears to be relatively more dependent on the visual code at this point.

It is possible that the present results could obtain if half the subjects used a visual code and half used a verbal code, interference effects from both pictorial and verbal suffixes then appearing in the pooled data. Examination of the individual scores shows that this is not the case. Although contaminated by order effects, individual data for the three terminal list positions shows that only 3 of the 18 subjects had a decrement in one of the interference conditions, but not in the other. The remaining 15 all showed decrements for both types of suffix as compared with the control condition.

Although the serial position curves, and the susceptibility of the recency effect to interference found in this experiment resemble those obtained in previous studies with auditory stimuli, it would be ill-judged to conclude that a *PVS*, or very-short-term visual store, underlies these results analogously to the *PAS* (the pre-categorical acoustic store) postulated by Crowder and Morton. The slow rate of presentation used here ensures that the interval between presentation and recall at any list position far exceeds any estimate of the duration of a visual sensory register. The results of this experiment should not therefore be taken as reflecting two stages in the transfer of material from a sensory register to a short term memory store, but rather the existence of two alternative codes within the short term store.

The rehearsal strategies employed by the subjects are interesting. Individual performance varied very greatly. The worst subject had a lowest score of only 7% correct in one condition, while the best subject achieved 100% correct in the same condition. The majority of subjects subvocalized the names of the pictured objects, but the better subjects reported using an associative stringing strategy. The pictured objects were related either spatially, or semantically by the construction of some kind of story context, and these methods yielded markedly better recall. The curiously improved performance evident at position 5 in both the picture prefix and the picture suffix conditions may result from some associative grouping strategy which is easier to apply to 12 pictures than to 10 pictures.

In effect, this experiment must be regarded as posing more problems than it solves. It suggests a number of further questions which need to be answered in order to confirm and extend the interpretation which has been offered. For example, what pattern of results emerges when the list items are all object names, and the redundant items either pictures, or further names? What is the effect of

shifting the rate of presentation, and of varying the method of recall? Can coding strategies be controlled and examined more formally? Further research along these lines now requires to be undertaken. This paper merely outlines the territory which needs to be explored in depth.

References

CROWDER, R. G. (1970). The role of one's own voice in immediate memory. *Cognitive Psychology*, **1**, 157–78.

CROWDER, R. G. and MORTON, J. (1969). Precategorical acoustic storage (PAS). *Perception and Psychophysics*, **5**, 365–73.

DALLETT, K. M. (1965). "Primary Memory:" the effects of redundancy upon digit recognition. *Psychonomic Science*, **3**, 237–8.

MORTON, J. (1968). Selective interference in immediate recall. *Psychonomic Science*, **12**, 75–6.

MORTON, J. and HOLLOWAY, C. M. (1970). On the absence of cross-modal suffix effects in short term memory. *Quarterly Journal of Experimental Psychology*, **22**, 167–76.

NEISSER, U., HOENIG, Y. and GOLDSTEIN, E. (1969). Perceptual organisation in the prefix effect. *Journal of Verbal Learning and Verbal Behaviour*, **8**, 424–9.

PAIVIO, A., ROGERS, T. B. and SMYTHE, P. C. (1968). Why are pictures easier to recall than words? *Psychonomic Science*, **11**, 137–8.

PARLETT, M. R. (1968). Systems of identification and storage of verbal stimuli. Unpublished doctoral thesis, University of Cambridge.

POTTER, M. C. and LEVY, E. I. (1969). Recognition memory for a rapid sequence of pictures. *Journal of Experimental Psychology*, **81**, 10–15.

PAUSE MEASURES DURING READING AND RECALL IN SERIAL LIST LEARNING

A. L. WILKES, P. LLOYD† AND I. SIMPSON

Department of Psychology, University of Dundee

An investigation is reported into pausing and grouping during the serial learning of letter strings, when presented randomly by length and when presented in ascending order. Mean pause times for the reading and recall of longer lists were significantly greater than for lists of shorter span due to extended pausing at specific list positions. In general, reading rhythms were duplicated during recall. Subjects were highly consistent in maintaining their level of pause duration across lists and responded to additions in list length by increasing their number of groups, not by increasing group size. Triadic sequences were the most popular form of spontaneous organization. Pausing measures are discussed as indices of organizational strategies.

Introduction

The tendency to chunk lists exceeding immediate memory span during acquisition is well established but the principles of organization that determine this pheno-menon are less clearly understood. It is extremely unlikely that there is a single ideal strategy for all forms of learning or, even within the same learning paradigm, that an optimal grouping pattern is insensitive to list content or the purpose recall will serve, e.g. Bower and Springston, 1970. The set of grouping strategies a subject can call upon is of theoretical and applied interest, as is any additional measure that can contribute to its analysis.

Belmont and Butterfield (1969) refer to the significance of using a direct experi-mental measure of acquisition. Such a measure would not deduce acquisition processes from recall measures and would, ideally, be sensitive to individual variations. One such measure they describe is a self paced task in which subjects were free to run through the items of a serial list at varying speeds of their own choice prior to the appearance of a probe item which then had to be correctly located in the list. The preferred intervals between list items during learning were recorded and constituted a record of the subject's characteristic hesitation pattern. It was found that subjects were consistent in their hesitation pattern from trial to trial and that the position of an item within a list, to a certain extent was predictive of hesitation duration: least time, for example, was spent pausing between terminal list items.

A related but potentially more general index of acquisition strategy is pausing behaviour during learning trials based on overt reading and recall. A number of recent studies of serial verbal learning support this generalization particularly those

† Now at University of Edinburgh.

that have manipulated the temporal order of items within a sequence and thus induced some experimentally predetermined pausing pattern; e.g. Ryan (1969); Neisser (1969); Bower and Winenz (1969); Winzenz and Bower (1970); Wilkes and Kennedy (1970). These experiments demonstrated the importance of grouping patterns for subsequent list learning, and retrieval, and established that grouping boundaries and extended pauses tend to concur. If temporal pauses in letter series are the major places at which recoding occurs, Bower and Springston (1970), then the monitoring of pausing behaviour during learning can serve as an index of subjective organization during the acquisition stage. The investigation of *spontaneous* pausing patterns during learning, as opposed to experimentally imposed patterns, has not however, been systematically studied. Suci (1967) investigated the learning of prose material structured so as to preserve or violate pause defined units and reported significant retardation in learning when such violation occurred. It was concluded that pause measures could serve to identify the transitions between organizational units. Thorpe and Rowland (1965) used pausing as one of a number of measures of spontaneous grouping patterns during serial list learning but relied upon subjective estimates of pause duration and thus do not provide a systematic or a continuous measure. The present investigation therefore set out to record spontaneous pausing behaviour when no prior structure had been imposed on the lists presented for serial learning. It was decided to monitor pausing at all positions within lists with the preliminary aim of establishing distributional and temporal data within a standard learning situation. Accordingly, a variety of different list lengths was presented randomly for learning in order to trace dependent changes in the pause distributions. In addition, the presentation of lists cumulatively increasing in length was investigated to permit some preliminary analysis of the organizational patterns indicated by the pause measures.

Experiment I

Method

Subjects

Ten female subjects, undergraduate students in the University of Dundee, participated in the experiment.

Materials

Ten different letter strings at each length from 3 to 11 items inclusively were prepared by random selection from the consonant set in the alphabet. Each of the resulting 90 letter strings was typed in upper case on cards with an equal spacing between items.

Procedure

All subjects learned a 3 item practice list and then one list at each of the remaining list lengths progressing randomly from length to length. The lists were presented visually using a tachistoscope and subjects read them aloud from left to right. Immediately a list had been read, it was replaced by a blank field and the subject tried to recite aloud the list items in their original order. This cycle of reading and recall was repeated until a criterion of three correct, consecutive recalls was reached. The interval between learning trials was 3 sec and this general procedure was repeated for all list lengths. It was arranged that no letters were repeated within a list and that no subject learned consecutive lists having letters in common. An interval of 15 sec separated the learning trials for different lists. The experimental

session took place in a sound deadened cubicle and lasted about 20 min. All sessions were recorded on tape.

Pause Measurement

The tapes for each subject were later processed using a voice key and pen recorder. The reading and recall sequences during the criterion trials were played back into the voice key and a resulting square wave output was then fed to a pen recorder which wrote a close approximation to each letter spoken. The paper speed used was 10 mm/sec and the distance between pen deflections was measured and taken as an estimate of pausing between list items.

Results and Discussion

An overall perspective on the changes in pausing that occurred with increasing list length is provided by the group means for each list. Table I includes the mean pause times for reading and recall separately and also the mean latencies that occurred between the completion of a reading of a list and the onset of recalling it, hereafter referred to as recall latency. All measures are based on averages from the three criterion trials.

TABLE I

Mean pause times for lists, reading and recall, and recall latency (msec). Experiment I

List	Reading	Recall	Recall latency
4	100	95	1314
5	143	135	1202
6	139	135	1356
7	153	133	1168
8	167	151	1308
9	167	200	1294
10	193	258	1381
11	201	251	1292

Newman-Keuls Comparisons.   ($P < 0.05$); reading: 11, 10, 9, 8 > 4; recall 11, 10 > 4 5 6 8; 9 > 4.

Generally, mean pause times increased for the longer lists, the effect being most pronounced during recall: list length, ($F_{7,63} = 6.49$; $P < 0.01$) and list length × mode of reproduction, ($F_{7,63} = 3.17$; $P < 0.01$). In contrast the recall latencies did not alter signficantly with list length, ($F_{7,63} = 1.17$, N.S.).

Considering any one list there was wide inter subject variation in reading and recall rates but subjects were consistent in maintaining their relative positions in mean pause times from list to list, both during reading and during recall. Kendall's Coefficient of Concordance, calculated across lists 4 to 11, indicated for reading, $W = 0.48$; χ^2 (9 df) = 34.8, $P < 0.001$; and for recall, $W = 0.35$; χ^2 (9 df) = 25.89, $P < 0.01$. Thus the increase in mean pause time with list length occurred within this consistent context of individual variation.

Such results could arise either from longer lists being read and recalled at a slow rate, pausing between all list items increasing, or, alternatively, increases in pause

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY

Volume 23   1971

Printed in Great Britain

Volume 23

Number 1, February 1971

NUMBER 2, MAY 1971

NUMBER 3, AUGUST 1971

NUMBER 4, NOVEMBER 1971

Author Index

duration could be restricted to particular intra list locations. Analysis of the pause ranges, list by list, revealed no tendency for the shortest pause times to increase during reading or recall whereas the longest pause durations did vary significantly: list length, $F(7,63) = 10 \cdot 17$; $P < 0 \cdot 001$; mode of reproduction, $F(1,9) = 6 \cdot 14$, $P < 0 \cdot 05$ and list length $\times$ mode of reproduction, $F(7,63 = 4 \cdot 09, P < 0 \cdot 001)$. As list length increased subjects responded by *selectively* extending their pause times during reading and recall although for the longer lists the extension was most marked during recall.

Further analysis of extended pauses, e.g. their frequency and location within lists, required a formal definition suitable for application at the level of individual subjects. It was decided such a definition should incorporate a subject's characteristic pausing level throughout a list as well as a comparative measure relating the pause time at a particular position to pause times at adjacent positions. These criteria were combined for a comparative classification of pause levels at all positions throughout a list as "Same", "Higher", or "Lower", than the pause level for the immediately preceding position. For every subject pausing at all but the first position within a list was categorized in this way by assigning pauses that differed in excess of 50% of the mean pause time for the whole list to the "Higher" or "Lower" categories, with intermediate levels being classified "Same". Having categorized every pause between list items in this way, extended pauses marking the boundaries of subjectively imposed groups were identified from "Higher/ Lower" and "Same/Lower" transitions. For the majority of the individual data, taken list by list, the application of the formal definition simply reproduced in an objective manner, grouping patterns that were clearly visible from the individual records. The mean group size at each list length during recall showed little variation, ranging from $2 \cdot 7$ to $3 \cdot 4$ list items. The majority of subjects responded to increases in list length by increasing the number of imposed groups, keeping group size on average around 3 items. A detailed analysis of grouping patterns, at the level of individual subjects revealed large inter subject and inter list variations. A second experiment is reported that had the aim of reducing subject variability thus permitting a qualitative analysis of the grouping patterns imposed.

Experiment II

In an attempt to reduce individual variations in grouping patterns it was decided to standardize the sequence of lists that subjects learned and to restrict the sets of letters from which specific list lengths were drawn.

Method

Twenty female subjects, undergraduate students in the University of Dundee, participated in the experiment. Nine sets of letters, varying in size from 3 to 11 items inclusively were compiled by randomly selecting the appropriate number of consonants from the alphabet. It was arranged that letter sets for consecutive list lengths did not contain any letter in common. From each set of letters, 20 different random combinations were constructed except in the case of the three letter set where the six possible combinations had to be repeated to provide the 20 letter strings required. Letter incidence was equalized across the sets used for each list length, as far as possible. The resulting 180 letter strings were typed in upper case on cards with equal spacing between items.

A subject progressed through the list lengths in the same, ascending order learning one list sequence at each length. Every subject experienced a different combination of the letters available for a specific list length. The presentation and learning procedures were the same as in Experiment I. During the pause measurement the paper speed used was 25 mm/sec.

Results and Discussion

As before, the first list of 3 items has been treated as a practice run and dropped from all subsequent analyses. Comparison with the pausing characteristics reported for Experiment I illustrated their generality: all main features were replicated despite the confounding effects of practice in the cumulative condition.

TABLE II
Frequency of structural groupings for each list. Recall. Experiment II

List	No structure	Doublet	Triplet	Quartet	Other
5	(9)	2:3 (5)	3:2 (6)	—	—
6	(4)	2:4 (4)	3:3 (9)	4:2 (3)	—
7†	(3)	2:5 (2)	3:4 (5) 3:2:2 (3)	4:3 (3)	5:2 (3)
8	(1)	2:2:2:2 (3) 2:2:4 (2) 2:3:3 (1) 2:6 (1)	3:3:2 (3) 3:2:3 (2) 3:5 (1)	4:4 (4) 4:2:2 (1)	6:2 (1)
9	—	2:2:3:2 (2) 2:2:5 (1)	3:3:3 (6) 3:6 (1) 3:2:4 (1) 3:4:2 (1)	4:5 (3) 4:3:2 (2)	6:3 (2) 5:4 (1)
10†	—	2:3:2:3 (1) 2:3:5 (1) 2:4:4 (1)	3:3:4 (5) 3:4:3 (1)	4:2:2:2 (2) 4:3:3 (1) 4:6 (1)	6:4 (3) 6:2:2 (1) 5:2:3 (1) 8:2 (1)
11†	—	2:2:4:3 (1) 2:3:3:3 (1)	3:3:3:2 (5) 3:3:5 (2) 3:4:4 (1) 3:3:2:3 (1)	4:2:2:3 (1) 4:2:3:2 (1) 4:3:4 (1) 4:4:3 (1) 4:5:2 (1)	5:2:4 (1) 5:3:3 (1) 7:4 (1)

† = Data for 1 subject missing.

Included in Table II are the grouping patterns recorded for each list with their frequency given in parentheses. The grouping sequences are classified (doublet, triplet, etc.) on the basis of the number of items in the first group. A triplet pattern was most popular at all lengths with the exception of list 8 where no clear preference was observed. Overall, of the 120 recorded instances of imposed groupings 45%

fell into the triplet category. The doublet and quartet categories occurred with equal but lower frequencies (21%) and an initial group size in excess of 4 items accounted for 13% of recorded cases. The most frequent specific grouping sequence at each list length is underlined in Table II and in all but one list these were drawn from the triplet category.

The grouping descriptions in Table II refer to the recall of lists only. An analysis of reading patterns in a similar manner indicated that the majority of subjects imposed their recall rhythm at the time of reading. The ascending presentation order made it possible to study the modification of grouping structure as list length increased. A Wilcoxon t-test indicated that in reading and recall the terminal group changed more frequently than the initial group in response to list length (reading, $N = 12$, $t = 11$, $P < 0.05$ two tailed; recall, $N = 13$, $t = 8$, $P < 0.01$ two-tailed). Associated with this tendency to incorporate the extra list item within the final group is the observation that the most popular grouping pattern at each list length also tended to follow a systematic cycle, disrupted at list 8, but taken up again for the remaining lists. The cycle consisted of repeated groups of 3 with remainder list items being incorporated into a final group up to a capacity of 4 items. Thus lists 6 and 9 elicited 3 : 3 and 3 : 3 : 3 as the most likely grouping structure and list 7 and 10 elicited 3 : 4 and 3 : 3 : 4. Such a sequence could occur if a terminal group incorporated extra list items until a capacity of 4 items was reached. Thereafter an extra triad could be generated and terminal group size vary cyclically between the limits of 2 and 4 items. Although the doublet and quartet categories do not permit similar generalizations the use of the terminal group as a container for the extra item would imply a rehearsal strategy similar to the differential treatment of items in primary and secondary memory store described by Belmont and Butterfield (1969). Obviously additional list items could not be indefinitely added to a terminal container group and a generative strategy would be an economic way of dealing with this problem. It should be noted that the method of identifying inter group pauses precluded the beginning and end item of a list from featuring as separate 1-item groups. This limiting condition is offset, however, by the observation that pause times at these locations tended to be brief and thus were unlikely under any definition to qualify as locations for an extended pause. The pausing records support other studies in presenting the triadic pattern as the most popular form of organization; in addition they seem to provide a sensitive means of mapping the strategic systems subjects can employ.

References

BELMONT, J. M. and BUTTERFIELD, E. C. (1969). The relations of short term memory to development and intelligence. In LIPSITT, L. P. and REESE, H. W. (Eds), *Advances in Child Development and Behaviour* Vol. 4. New York: Academic Press.

BOWER, G. H. and WINZENZ, D. (1969). Group structure, coding and memory for digit series. *Journal of Experimental Psychology Monographs*, 80, No. 2, Part 2, 1-17.

BOWER, G. H. and SPRINGSTON, F. (1970). Pauses as recoding points in letter series. *Journal of Experimental Psychology*, 83, 421-30.

NEISSER, U. (1969). The role of rhythm in active verbal memory: serial intrusions. *American Journal of Psychology*, **82**, 540–6.

RYAN, J. (1969). Grouping and short term memory: different means and patterns of grouping. *Quarterly Journal of Experimental Psychology*, **21**, 137–47.

SUCI, G. J. (1967). The validity of pause as an index of units in language. *Journal of Verbal Learning and Verbal Behaviour*, **6**, 26–32.

THORPE, C. E. and ROWLAND, G. E. (1965). The effect of "natural" grouping of numerals on short term memory. *Human Factors*, **7**, 38–44.

WILKES, A. L. and KENNEDY, R. A. (1970). The relative accessibility of list items within different pause defined groups. *Journal of Verbal Learning and Verbal Behaviour*, **9**, 197–202.

WINZENZ, D. and BOWER, G. H. (1970). Subject-imposed coding and memory for digit series. *Journal of Experimental Psychology*, **83**, 52–6.

THE EFFECTS OF ACOUSTIC AND SEMANTIC SIMILARITY ON CONCEPT IDENTIFICATION

VERONIKA COLTHEART†

Monash University

Since both acoustic and semantic similarity exert an influence on memory, the role of memory in concept identification (CI) was investigated by varying the acoustic and semantic similarity of the stimuli used in the CI task. Varying acoustic similarity had no effect on CI, but CI was significantly impaired when the dimensions of a CI task were semantically similar.

In recent years various suggestions have been made concerning the role of memory in simple concept identification (CI). A sampling-with-replacement assumption (implying no memory) was originally proposed in models of CI (Bower and Trabasso, 1963, 1964; Restle, 1962), but was not supported by subsequent experimental evidence (Holstein and Premack, 1965; Levine, 1962, 1966; Merryman, Kaufmann, Brown and Dames, 1968; Trabasso and Bower, 1966) and henceforth it was recognized that memory is used in concept identification.

Trabasso and Bower (1966) suggested that subjects remember earlier stimuli and the responses corresponding to them. Restle and Emmerich (1966) gave subjects varying numbers of CI tasks concurrently and obtained evidence for the importance of memory. They concluded that subjects encode descriptions of the stimuli and feedback, maintaining these in short-term memory with the aid of rehearsal. However, direct investigations of memory for stimuli and responses using both recognition and recall procedures, indicate that such memory is minimal (Bourne and O'Banion, 1969; Calfee, 1970; Coltheart, 1971*b*).

Levine (1966) proposed that subjects monitor a pool of possible hypotheses using successive trials to reduce this pool until only one hypothesis, the correct one, remains. Hence, subjects covertly test all hypotheses simultaneously although overt responses to the stimuli are based on only one hypothesis at any time. In a binary-valued *n*-dimensional task, subjects begin with a set of $2n$ hypotheses and this set is reduced to *n* hypotheses by the first stimulus and feedback trial. One of these *n* hypotheses is then used as the basis for the response given on the next trial. The second and subsequent trials, provided that feedback is given, enable subjects to reduce the hypothesis set by degrees directly dependent on the particular stimulus sequence presented. According to Levine, subjects have to perform the following tasks: they must remember their hypothesis set from trial to trial, perform logical operations on this set in accordance with the

† Now at Psychology Department, University of Guelph, Guelph, Ontario, Canada.

information provided by each feedback trial, and they must also respond (with a binary choice response) to each stimulus that is presented. For Levine, memory is involved because hypotheses are encoded in the form of a verbal description and retained by rehearsal.

Chumbley (1967, 1969) formalized Levine's (1966) theory, deriving predictions both for it and for Trabasso and Bower's (1966) theory. These predictions were tested in an experiment which varied both the number of tasks performed concurrently and the length of the post-feedback interval. Chumbley concluded that a theory which incorporated memory for hypotheses was better able to explain not only his results but also those of Levine (1966) and Richter (1965), than a theory which states that subjects remember stimulus–response information. Other studies have found evidence of memory for hypotheses (Coltheart, 1971b; Erickson, 1968; Levine, 1969). In particular, Erickson (1968) suggested that recently tested hypotheses are stored in short-term memory until new information pushes earlier information out.

Levine (1969) presented evidence showing that, at any stage during CI, after an error, the hypothesis sampled is more likely to be consistent with information from the first and the most recent feedback trials than it is with the intervening feedback trials. He attributes the recency effect to storage in short-term memory and the primacy effect to a longer-term memory.

Clearly, there is considerable evidence to suggest that there is memory in some form for hypotheses in concept identification. Before proceeding further it is necessary to consider the meaning of the term, "hypothesis". An hypothesis is usually understood to refer to a rule based on one of the dimensions of a simple CI task. Since the theories and research discussed have concerned only tasks with binary-valued dimensions, in an n-dimensional task there are exactly $2n$ possible hypotheses, two for each dimension. There is evidence to suggest that, when a subject makes an error while testing a hypothesis based on a particular dimension, he rejects not just that particular hypothesis, but also the complementary hypothesis based on that dimension (Levine, 1969; Levine, Miller and Steinmeyer, 1967). Others have suggested that subjects deal with whole dimensions rather than hypotheses (Kendler and Kendler, 1962; Glanzer, Huttenlocher and Clark, 1963; Tighe and Tighe, 1968). Consequently, an adequate theory of memory in CI may need to be formulated in terms of memory for dimensions rather than for hypotheses despite the fact that responses during CI tend to be based on specific hypotheses.

Various suggestions have been made about the kind of memory storage involved in remembering hypotheses: short-term memory and possibly a longer-term memory also. However, no direct evidence regarding the nature of memory storage used in concept identification has been obtained. The experiments to be described were aimed at investigating memory in concept identification by attempting to relate performance on these tasks to current research in memory.

In a series of papers Conrad (1959, 1962, 1964; Conrad and Hull, 1964) has reported evidence that short-term memory uses an acoustic (or articulatory) code. The essential finding is that when visually presented letters are erroneously recalled, the substitution errors tend to be letters that *sound* like the original item

(Conrad, 1964). In fact, the memory errors could be predicted from listening errors subjects made when trying to identify single letters spoken against a background of white noise. Baddeley (1966a, 1966b) reported differential effects of acoustic and semantic similarity on memory. Performance on a list of acoustically similar words was significantly inferior to performance on a comparable control list for the first three trials, but thereafter it reached the level for the control list (Baddeley, 1966a). The opposite tendency was observed with a list of semantically similar items. Performance on this list improved at a rate comparable to control list performance but reached asymptote by the third trial whereas control list performance continued to improve and reached a much higher asymptote. Baddeley concluded that while short-term memory seems to use an acoustic code, long-term memory depends mainly on a semantic code. The investigation of the tip-of-the-tongue phenomenon (Brown and McNeill, 1966) also supports the existence of acoustic and semantic encoding in memory as does the fact that the build-up and release of proactive interference can be achieved by appropriate manipulation of acoustic similarity (Coltheart and Geffen, 1970).

The effects of acoustic and semantic similarity were investigated in the context of CI. Since subjects seem to rely to some extent on memory for hypotheses (or dimensions) and if, as some have suggested, a short-term storage is used, memory should be adversely affected if the names of the dimensions are acoustically similar. Therefore, performance on a task with acoustically similar dimensions might be inferior to performance on a comparable control task.

However, since acoustic similarity has adverse effects in such a large variety of short-term memory tasks, it is conceivable that it may even prevent subjects from remembering their current hypothesis from one trial to the next. For this reason, it seemed advisable to construct a task in which only the values of the dimensions were acoustically similar. If the dimensions are not themselves acoustically similar, the subject need not have difficulty in remembering which dimensions he has tested, or which remain in his hypothesis pool, but he may find it difficult to remember which of the two possible hypotheses based on a particular dimension, he is currently testing. Any such difficulty could result in misapplication of trial-by-trial feedback information.

Since Erickson, Zajkowski and Ehmann (1966) have shown that response latencies change in an orderly way during CI, it was decided that latencies should be recorded and systematic comparisons made among subjects performing different CI tasks.

Experiment I

The first experiment compared performance on two tasks which had identical acoustically dissimilar dimensions but whose values differed. One task had acoustically similar values for each dimension, the other had not. The two letters most frequently confused with one another, F and S, and the two letters least likely to be confused, B and X, were chosen from Conrad's (1964) data as the values for each dimension in the two tasks. In both tasks the dimensions were four different colours. A simple rule using the colour yellow as the dimension would be: if the yellow letter on the slide is a B, "left" (the response) is correct; if

the yellow letter is an X, "right" is correct, for the control task. For the experimental task the corresponding rule would be: if the yellow letter on the slide is an F, "left" is correct and if the yellow letter is an S, "right" is correct.

Method

Subjects

Sixty undergraduate students at Monash University acted as subjects and were paid 50 cents for doing so. None had participated in a CI experiment previously. Half the subjects were assigned randomly to Group BX and the other half to Group FS.

Stimulus materials

Two practice problems, one two-dimensional and one three-dimensional, were constructed. The two dimensions of the first practice task were the colours, pink and purple, and the values were the shapes, square and triangle. The second practice task had the colours green, brown and orange as dimensions, and the digits, 1 and 2 as values. The three-dimensional task and the four-dimensional main problems were photographed on

BLACK   RED
BLUE   YELLOW
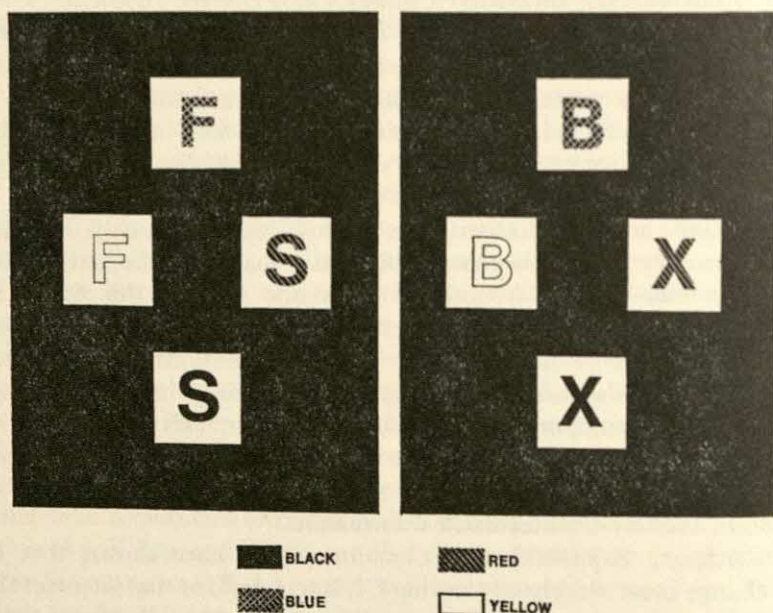
FIGURE 1. A pair of matching stimuli selected from the experimental and control conditions of Experiment I.

Kodachrome slides. The four-dimensional main problems both used the dimensions red, yellow, blue and black and the values were F and S for one task and B and X for the other. For all tasks, the positions of the coloured figures on each stimulus varied randomly. The main problems each contained four coloured letters and these were located at the top, bottom, left and right of the slide. The position of a letter of a particular colour was varied randomly except that care was taken to construct two tasks that were identical except for the fact that where there would be a B on one there was always an F on the other in the same colour and where an X would appear, the other task would have a corresponding S. Figure 1 presents an example. No colour appeared in one of the four positions on the slide more than four times out of the total of 16 stimuli constructed.

Apparatus

A Kodak carousel projector showed the slides on a screen in front of the subject who sat in a chair equipped with an armrest and a response switch which could be pushed to the left or to the right. The projector was connected to a series of timers so that the experimenter could record the time between the onset of a stimulus and the subject's response. The subject's response, left or right, automatically activated the appropriate feedback for that trial. The feedback light and slide remained on view for another 4 sec and then the next slide came on. Feedback was provided by means of two green lights located at the lower left and right hand corners of the screen. The left hand light was illuminated on trials for which the response "left" was correct and the right hand light on trials for which the response "right" was correct.

Procedure

When the subject was seated, the experimenter described the dimensions and values of the two-dimensional practice problem, and demonstrated the various ways of sorting the cards into two piles. The dimensions and values of the three-dimensional task and the type of rule were described along with the way responses were to be made and the operation of the feedback lights. Responding was self-paced. The subject's response switched on the response feedback light and activated a timer so that feedback and stimulus remained on view for 4 sec after the response before the next stimulus was automatically presented. The subject had to identify the concept arbitrarily chosen by the experimenter and was required to respond until eight consecutive correct responses had been completed. Similar instructions were given for the four-dimensional task and the subject was told that the criterion to be met was 16 consecutive correct responses. For all three tasks, the instructions stressed that the location of the dimensions on the stimuli varied randomly and was irrelevant for the solution.

Both groups BX and FS were given identical instructions and treatment except for the fact that the values of the last task differed for the two groups. The stimulus sequence and feedback were identical for the two tasks so that the trial-by-trial information provided to both groups was equivalent.

Results

The mean trial of last error (TLE) and standard deviations were 7·03 and 5·76 for Group BX and 8·57 and 7·58 for Group FS, respectively. The difference between the groups was not significant ($t = 0·87$; $df = 58$; $P > 0·30$). Latencies of presolution responses were classified as after-error (the previous response was an error) or after-correct (the previous response was correct). For every subject who had at least one after-error and one after-correct latency a median after-correct and a median after-error latency were computed and a t-test was performed for each task. For both tasks the mean after-error (AE) latency was significantly greater than the mean after-correct (AC) latency. These results are presented in Table I. There were no differences between latencies for Group BX and Group FS.

Discussion

Since no differences could be detected between performance on a task with acoustically similar values and performance on a control task, it was concluded that subjects do not have difficulty in remembering their current hypothesis from one trial to the next when the dimension values are acoustically similar. One subject did report difficulty in remembering whether "S" or "F" was assigned to

TABLE I

Presolution latency data for Experiment I

Group	Mean after-error (sec)	S.D.	Mean after-correct (sec)	S.D.	t	P
BX	5·14	3·48	3·73	2·24	2·08	<0·05
FS	4·71	2·24	3·19	1·44	2·73	<0·01

the left, and said she resolved this by recalling: "S = socialist = left and F = fascist = right".

Experiment II

This experiment was designed to investigate whether CI is impeded if the dimensions are acoustically similar. Again, Conrad's (1964) results were used to select four acoustically similar letters to serve as dimensions for the experimental task and four acoustically dissimilar letters as dimensions for the control task. The same two dissimilar colours were the values for all of the dimensions in both tasks. Hence each stimulus was composed of four letters which could appear in one of two colours, red or blue. A simple concept rule using the letter B as the relevant dimension would be: if the B on the slide is blue, the correct response is "left"; if the B is red, the response is "right".

Method

Subjects

A further 60 Monash undergraduates acted as subjects for 50 cents payment. Thirty were randomly assigned to the experimental group and 30 to the control group.

Stimulus materials

Stimuli for a three-dimensional and two four-dimensional tasks were photographed and made into Kodachrome colour slides. The three dimensions of the practice task were the letters, F, D, and Q, and their values were the colours, red and blue. The values for both four-dimensional tasks were the colours red and blue. The four dimensions of the experimental task were the letters B, P, V and C and the letters B, S, Y and J for the control task. All four letters appeared on every slide but letter location, top, bottom, left, or right, varied randomly from slide to slide. The two main tasks were constructed so that the letter B appeared in the same position and colours in both. The location and colour of the letter C in the experimental task was matched by J in the control task. Similarly, the letters P and V in the experimental task were matched by S and Y respectively, in the control task. Figure 2 shows a stimulus from each task. Since the stimuli for both tasks were constructed in the way described, it was possible to present stimulus and response feedback sequences that provided logically equivalent trial-by-trial information to both groups of subjects.

Apparatus and procedure

The apparatus described earlier was used and the procedure followed was similar to that of Experiment I except that only one practice task preceded the main task. Again, instructions emphasized that letter location varied randomly and was irrelevant to solution.

Results

The mean TLE and standard deviation for the task with acoustically similar dimensions were 7·13 and 8·61, and the mean TLE and standard deviation for the control task were 6·03 and 4·21. The difference between the groups was not significant, $t = 0.62$; $df = 58$; $P > 0.50$.
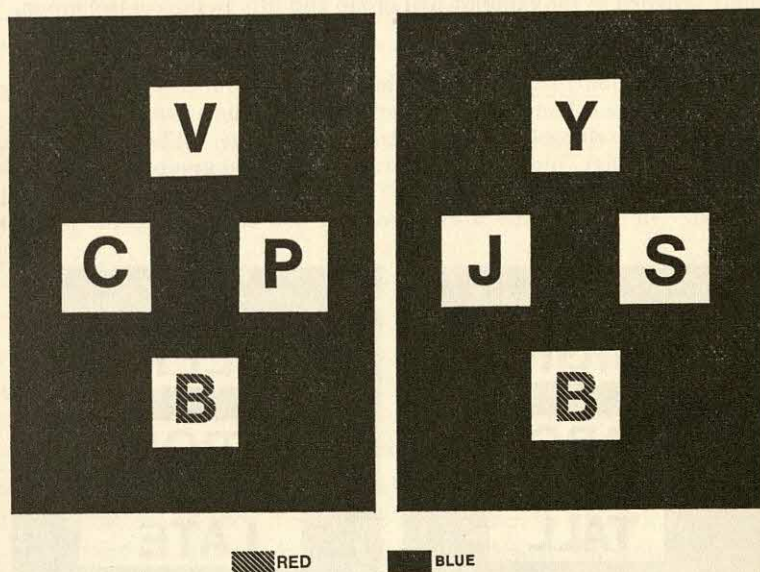
FIGURE 2. A pair of matching stimuli selected from the experimental and control conditions of Experiment II.

Again, AC latencies were significantly shorter than AE latencies for both tasks while latencies obtained in the control and experimental tasks did not differ. The results are presented in Table II.

TABLE II

Presolution latency data for Experiment II

Group	Mean after-error (sec)	S.D.	Mean after-correct (sec)	S.D.	t	P
Acoustically similar	3·71	1·78	2·41	0·71	3·93	<0·005
Control	3·74	1·75	2·67	1·35	2·60	<0·01

Experiment III

In Experiment III performance on a task with semantically similar dimensions was compared to performance on a control task. Four of the adjectives used by

Baddeley (1966) in his semantically similar list and four from the control list were chosen as the dimensions for the two tasks.

Method

Subjects

A further 100 Monash undergraduates acted as subjects for 50 cents payment. Fifty were randomly assigned to the experimental group and fifty to the control group.

Stimulus materials

Stimuli for a three-dimensional practice and two four-dimensional tasks were photographed on Kodachrome colour slides. The practice task dimensions were the words, hot, thin, and safe, and the values were the colours, red and blue. The main task dimensions were the words, high, large, big and tall for the experimental group and the words, strong, late, deep and old for the control group. The values were the colours, red and blue for every dimension. In all tasks the words appeared in a column and the position of each

HIGH	DEEP
LARGE	STRONG
TALL	LATE
BIG	OLD

▨ RED   ■ BLUE
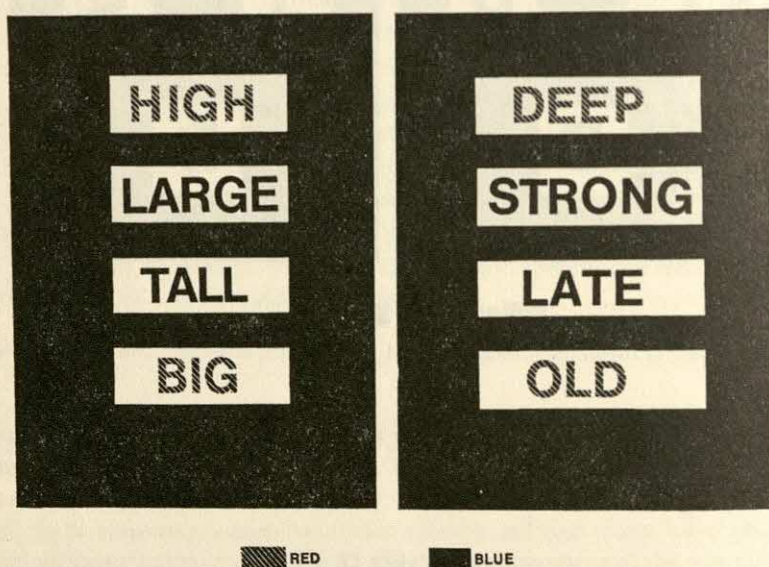
FIGURE 3. A pair of matching stimuli selected from the experimental and control conditions of Experiment III.

word varied randomly from slide to slide. Figure 3 provides a stimulus from the experimental and control groups. The colour and location of the word, "high", in the experimental task was matched by the word, "deep", in the control task. Similarly, "large" was matched by "strong", "big" by "old" and "tall" by "late" in the control task. Again, the stimulus and response feedback sequences presented to subjects in both groups were equivalent in terms of the trial-by-trial information they provided. Thus, a rule based on the dimension, high, would be: if the word, high, is red, "left" is correct; if the word, high, is blue, "right" is correct. For the control task, the corresponding rule would be based on "deep" and would be: if the word, deep, is red, "left" is correct; if the word, deep, is blue, "right" is correct.

Apparatus and procedure

These were identical to apparatus and procedure used in Experiment II except that the dimensions of the tasks were different.

Results

The mean TLE and standard deviation for the experimental group were 9·68 and 8·61, and for the control group they were 6·50 and 5·11, respectively. The difference between the two groups was significant ($t = 2.22$; $df = 98$; $P < 0.05$).

TABLE III

Presolution latency data for Experiment III

Group	Mean after-error (sec)	S.D.	Mean after-correct (sec)	S.D.	t	P
Semantically similar	3·11	1·28	2·27	0·53	4·50	<0·005
Control	3·25	1·55	2·17	0·49	4·53	<0·005

Table III presents AC and AE latencies for both tasks. AC latencies were shorter than AE latencies in both tasks. Latencies did not differ between tasks.

Discussion

Although the subjects presented with CI tasks having either acoustically similar values or dimensions do not find these more difficult than comparable control tasks, it seems that CI is more difficult if the dimensions are semantically similar.† It should be noted that these results were obtained when subjects perform under the usual conditions which involve self-paced responding and stimulus and feedback availability for several seconds after each response. Under conditions providing greater memory demands or less information-processing time different results might have been obtained. Since subjects needed significantly more trials to complete a CI task with semantically similar dimensions, it is surprising that their latencies did not differ from the latencies of the control group. Perhaps response latency is a less sensitive index of the CI process than formerly supposed (Erickson *et al.*, 1966).

Since CI performance is not affected if the dimensions are acoustically similar and since it is significantly hampered if the dimensions are semantically similar, it is concluded that the information subjects attempt to remember whilst performing a CI task is stored primarily in a semantic rather than an acoustic code. If the view of Baddeley (1966*b*), that acoustic encoding is characteristic of short-term memory while semantic encoding is used by long-term memory, is correct, then it would seem that memory for hypotheses (or dimensions) in CI involves the long-term store. However, a problem for Baddeley's proposal is that there is evidence to suggest that acoustic encoding exists in long-term memory as well (Brown and McNeill, 1966; Bruce and Murdock, 1968; Dallett, 1966). Perhaps the resolution of the conflicting evidence on this issue is provided by Bruce and

---

† Calculations indicate that the *t*-tests used to analyse the data of these experiments are of approximately equivalent power; the larger N in Experiment 3 is compensated for by the greater variability.

Crowley's (1970) study which shows that acoustic coding can occur in long-term memory but only in rather exceptional circumstances, suggesting that the semantic code is in fact the dominant one in long-term memory.

The possibility that subjects use long-term memory during CI derives support elsewhere (Coltheart, 1971a). Concept identification is unimpaired when three uninformative trials intervene between every trial on which feedback is provided. The interposition of three such trials between successive feedback trials considerably increases the short-term memory demands of the task without any noticeable effect on performance, suggesting that whatever subjects remember during CI is relatively unaffected by the passage of time.

References

BADDELEY, A. D. (1966a). The influence of acoustic and semantic similarity on long-term memory for word sequences. *Quarterly Journal of Experimental Psychology*, **18**, 302–9.

BADDELEY, A. D. (1966b). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Quarterly Journal of Experimental Psychology*, **18**, 362–5.

BOURNE, L. E., JR. and O'BANION, K. (1969). Memory for individual events in concept identification. *Psychonomic Science*, **16**, 101–3.

BOWER, G. and TRABASSO, T. (1963). Reversals prior to solution in concept identification. *Journal of Experimental Psychology*, **66**, 409–18.

BOWER, G. and TRABASSO, T. (1964). Concept identification. In ATKINSON, R. C. (Ed.), *Studies in Mathematical Psychology*. Pp. 32–94. Stanford: Stanford University Press.

BROWN, R. and McNEILL, D. (1966). The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior*, **5**, 325–37.

BRUCE, D. and CROWLEY, J. J. (1970). Acoustic similarity effects on retrieval from secondary memory. *Journal of Verbal Learning and Verbal Behavior*, **9**, 190–6.

BRUCE, D. and MURDOCK, B. B., JR. (1968). Acoustic similarity effects on memory for paired associates. *Journal of Verbal Learning and Verbal Behavior*, **7**, 627–31.

CALFEE, R. C. (1969). Recall and recognition memory in concept identification. *Journal of Experimental Psychology*, **81**, 436–40.

CHUMBLEY, J. I. (1967). The memorization and manipulation of sets of hypotheses in concept learning. Unpublished doctoral dissertation, Indiana University.

CHUMBLEY, J. I. (1969). Hypothesis memory in concept learning. *Journal of Mathematical Psychology*, **6**, 528–40.

COLTHEART, M. and GEFFEN, G. (1970). Grammar and memory: I. Phonological similarity and proactive interference. *Cognitive Psychology*, **1**, 215–24.

COLTHEART, V. (1971a). Memory in concept identification. Ph.D. thesis, Monash University.

COLTHEART, V. (1971b). Memory for stimuli and memory for hypotheses in concept identification. *Journal of Experimental Psychology*, **89**, 102–8.

CONRAD, R. (1959). Errors of immediate memory. *British Journal of Psychology*, **50**, 349–59.

CONRAD, R. (1962). An association between memory errors and errors due to acoustic masking of speech. *Nature*, **193**, 1314–5.

CONRAD, R. (1964). Acoustic confusions in immediate memory. *British Journal of Psychology*, **55**, 75–84.

CONRAD, R. and HULL, A. J. (1964). Information, acoustic confusion and memory span. *British Journal of Psychology*, **55**, 429–32.

DALLETT, K. M. (1966). Effects of within-list and between-list acoustic similarity on the learning and retention of paired associates. *Journal of Experimental Psychology*, **72**, 667–77.

ERICKSON, J. R. (1968). Hypothesis sampling in concept identification. *Journal of Experimental Psychology*, **76**, 12–18.

ERICKSON, J. R., ZAJKOWSKI, M. M. and EHMANN, E. D. (1966). All-or-none assumptions in concept identification: Analysis of latency data. *Journal of Experimental Psychology*, **72**, 690–7.

GLANZER, M., HUTTENLOCHER, J. and CLARK, W. (1963). Systematic operations in solving concept problems: A parametric study of a class of problems. *Psychological Monographs*, **77** (1, Whole No. 564).

HOLSTEIN, S. B. and PREMACK, D. (1965). On the different effects of random reinforcement and presolution reversal on human concept identification. *Journal of Experimental Psychology*, **70**, 335–7.

KENDLER, H. H. and KENDLER, T. S. (1962). Vertical and horizontal processes in problem solving. *Psychological Review*, **69**, 1–16.

LEVINE, M. (1962). Cue neutralization: the effects of random reinforcements upon discrimination learning. *Journal of Experimental Psychology*, **63**, 438–43.

LEVINE, M. (1966). Hypothesis behaviour by humans during discrimination learning. *Journal of Experimental Psychology*, **71**, 331–8.

LEVINE, M. (1969). Neo-noncontinuity theory. In BOWER, G. H. and SPENCE, J. T. (Eds), *The Psychology of Learning and Motivation*, Vol. 3. Pp. 101–34. New York: Academic Press.

LEVINE, M., MILLER, P. and STEINMEYER, C. H. (1967). The none-to-all theorem of human discrimination learning. *Journal of Experimental Psychology*, **73**, 568–73.

MERRYMAN, C., KAUFMANN, B., BROWN, E. and DAMES, J. (1968). The effects of "rights" and "wrongs" on concept identification. *Journal of Experimental Psychology*, **76**, 116–9.

RESTLE, F. (1962). The selection of strategies in cue learning. *Psychological Review*, **69**, 329–43.

RESTLE, F. and EMMERICH, D. (1966). Memory in concept attainment: effects of giving several problems concurrently. *Journal of Experimental Psychology*, **71**, 794–9.

RICHTER, M. L. (1965). Memory, choice and stimulus sequence in human discrimination learning. Unpublished Ph.D. dissertation, Indiana University.

TIGHE, T. J. and TIGHE, L. S. (1968). Perceptual learning in the discrimination processes of children: An analysis of nine variables in perceptual pretraining. *Journal of Experimental Psychology*, **77**, 125–34.

TRABASSO, T. and BOWER, G. (1964). Memory in concept identification. *Psychonomic Science*, **1**, 133–4.

TRABASSO, T. and BOWER, G. (1966). Presolution dimensional shifts in concept identification: a test of the sampling with replacement axiom in all-or-none models. *Journal of Mathematical Psychology*, **3**, 163–73.

THE LEARNING OF SURFACE STRUCTURE CLUES TO DEEP STRUCTURE BY A PUPPET SHOW TECHNIQUE

RICHARD F. CROMER

MRC Developmental Psychology Unit, London

Fifty-six children were classified in terms of their developmental level on the linguistic structure "John is easy/eager to please", and an experiment was carried out teaching them "new" (nonsense) words in related transformations by a puppet show technique. These transformations serve as frames for differentiating those new words which indicate a surface structure derived from a transformed deep structure, from those which do not. Results showed that children used different strategies at different developmental levels. Of children who did not use fixed strategies, only children who had correctly performed on all normal adult words ("Passers") were able to learn which deep structures were called for by some of the new words; all other groups performed at chance levels on all types of new words. However, even these Passers performed correctly only on words given in structures indicating that the tested structure was to be treated as a transformation. The possibility of a language-learning principle related to the marked/unmarked distinction in natural languages is discussed.

Introduction

There is increasing evidence that traditional learning theory principles such as imitation, frequency, practice and reinforcement, though important in explaining some aspects of language learning, are by themselves inadequate to account for the observed linguistic changes in children acquiring language (see, e.g. Bellugi, 1967; Bellugi-Klima, 1969; Brown, Cazden and Bellugi-Klima, 1969). Furthermore, there has been increasing interest in linguistic universals (see, e.g. Bach and Harms, 1968; Chomsky, 1965; and Greenberg, 1966a, 1966b), and it has been suggested that some of these may be applicable to the language acquisition process itself (Jakobson, 1968; McNeill, 1970; Miller and McNeill, 1969). One such universal, the "marked"/"unmarked" concept, has been extensively reviewed by Greenberg (1966a).

Surveying languages of the world, Greenberg has found the marked/unmarked contrast to have a high degree of generality in that it is applicable to phonological, grammatical and semantic aspects of language. A simple example using lay rather than linguistic terminology will serve to illustrate this notion. In English, the general way to form a plural is to add -s or -es to the singular. Thus, the plural ("lads", "lasses", etc.) is said to be "marked" in contrast to the singular ("lad", "lass", etc.) which is unmarked. Marking is considered to add something—some meaning or information—to the unmarked term. Sometimes this additional information acts to restrict or limit the term to particular meanings

(see Joos, 1964). This principle may be of special importance as it has been observed that in many linguistic contrasts the member which is marked is universally the same in all languages where that contrast is made. Thus, for example, in the phonological contrasts voiced/unvoiced, aspirated/unaspirated and nasal/non-nasal, it is the first number of each pair that is marked. And in grammar, it is always the plural, not the singular, and the passive, not the active, which is marked. It should be noted that in some languages, both members of a particular contrast pair may be differentially marked, but in languages where only one member is marked, it is universally the same side of the contrast to which marking is applied.

The marked/unmarked distinction has been extended to account for a number of linguistic contrasts, some of which do not have elements such as -s or -es to mark them overtly. For example, Greenberg (1966a) gives evidence for an hypothesized semantic marking in his discussion of semantic universals based on categories of marked and unmarked adjectival opposites, and in these instances there is no special linguistic element or inflection to indicate which is the marked form. Thus, if one asks "How deep is the water?" "deep" is considered unmarked since the water may be deep or shallow. But when one asks "How shallow is the water?", the adjective, "shallow", is seen as marked since it is already implied that the water is shallow, and one is simply asking how much so. In this case, there is no linguistic element attached to the word to show that it is marked.

Miller and McNeill (1969) have suggested, based on Greenberg's findings, that unmarked features are the ones produced as a matter of course, i.e. ones not requiring special decisions on the speaker's part. Linking this to language acquisition, they have proposed that the unmarked features are those which are unspecified in the base structure of sentences. They argue that since some evidence indicates that the first grammatical efforts of children are confined to the base structure, and as there is no evidence that children at that stage use transformational rules, children should first acquire the universally unmarked features of language, and only later begin to use marked forms. This has been found semantically, for example, where many young children treat the marked form "less" as if it indicated the unmarked form "more" (Donaldson and Balfour, 1968).

In order to study this principle and the part it may play in the child's acquisition of language at the grammatical level, it is possible to use linguistic structures which are acquired at a relatively late age in the course of acquisition. Carol Chomsky (1969) has studied several grammatical structures which are acquired between the ages of five and ten. One of them, of the type "John is eager to please/John is easy to please" is especially interesting for what it may reveal. Take the structure: "John is ———— to please". In "John is eager to please", it is said that the deep structure subject and the surface structure subject coincide, and it is John who is the actor, i.e. it is John who does the pleasing. But in "John is easy to please", adult English speakers somehow know that someone other than John does the pleasing. The surface or superficial subject "John" is not the deep subject, i.e. the surface structure is transformed from a different deep structure. It should be noted that the ordering of the sequence of classes of the two sentences is identical. The only linguistic cue to the recovery of the proper deep structure (in order to interpret the sentence) is the particular adjective which is used.

Adjectives can be classified into many different groups based on their privileges of occurrence (e.g. see Householder, Alexander and Matthews, 1964). For purposes of clarity in the present experiment, however, they are classified into three types as regards this linguistic structure. Adjectives like "eager" which indicate that the surface subject and deep structure subject coincide, have been labelled S-adjectives. A second class of adjectives, which like "easy" indicate that someone other than the surface subject is the actor, are designated O-adjectives. Finally, there is a class of adjectives which are ambiguous in this structure. When these are used, the interpretation of the actor or deep subject depends on the context, situation, and other cues. These are called A-adjectives.

It may be possible to consider the S- and O-adjective types as "marked" classes since they both provide special additional information as to sentence structure in contrast to the A-adjectives. S-adjectives indicate that one must recover a deep structure which has not had any basic relationships transformed, whereas O-adjectives demand recovery of a deep structure in which the basic relationships have been transformed, making an object appear to be a subject in the surface structure.

It would be interesting to know how these adjectives acquire their information value. One possibility is that they are marked by being heard in related transformations. The particular transformations which would be significant for such learning can be called "differentiating" frames, i.e. linguistic frames which differentiate between S- and O-type adjectives in terms of whether these adjectives are grammatically acceptable in the frame. For example, the frame "He was ———— to praise her" allows only S-type (or A-type) adjectives, but excludes the use of O-type adjectives. Thus one may say "He was *eager* to praise her", but not "He was *easy* to praise her". Conversely, in a different frame, one may say "For him to praise her was *easy*", but the sentence "For him to praise her was *eager*" is not grammatical. This second frame, then, accepts O-type (or A-type) adjectives, but rejects S-type adjectives. Hearing a "new" adjective in structures where S-adjectives can appear but O-adjectives cannot, or where O-adjectives can appear but S-adjectives cannot, may be the manner by which new adjectives are effectively "marked" as indicating whether a surface structure is or is not the result of particular transformations.

According to the Miller and McNeill hypothesis, mentioned earlier, young children should interpret all sentences of this structure on the basis of an untransformed deep structure. Thus, although both S- and O-adjectives are said to be differentially marked in contrast to the A-adjectives, children who have not yet acquired the rules for their use should interpret all sentences as if S-adjectives were being used. By using words in the child's own vocabulary in this structure, it has been shown that young children do indeed interpret sentences in this way (Chomsky, C., 1969; Kessel, 1970). They use what can be called the "primitive rule" of always assigning actor or deep subject status to the surface subject regardless of what adjective is used in the space. That this is specific to this structure is shown by the fact that they can interpret passive sentences correctly. Cross-sectional developmental studies (Cromer, 1970) show that at a later age (about mental age 6:3 on some tests), the child abandons this primitive rule and gives a

mixed interpretation when presented with several instances of this structure—some of these coinciding with adult interpretation and some not. Finally (often as late as chronological age 9:0 or 10:0), the child reaches a point where he interprets correctly all common uses of this structure. There are, then, two changes to be accounted for in the explanation of the manner by which the child acquires adult competence on sentences of this type. The first has to do with linguistic rule change, and the second with those specific changes of interpretation of particular lexical items necessary to attain agreement with adult interpretation.

The purpose of this experiment is to study this second change, in which the child must "learn" which surface features of language input (in this case, the particular adjectives) indicate which deep or interpretational meanings are called for. The intent is to examine, more systematically, a finding indicated in the final part of an experiment recently reported (Cromer, 1970) on the manner of children's acquisition of classes of adjectives by hearing "new" words in differentiating transformational frames.

Specifically, the aims of this experiment will be to see (1) whether children can learn the structure or "cue value" of "new" (nonsense) adjectives presented in differentiating frames in a simulated natural language situation, (2) whether groups of children at various developmental stages (as regards this structure) use differing strategies in the learning task and whether these strategy differences are quantifiable in terms of differing response latencies, (3) whether children at the various stages evidence differential ability to benefit from these frames, and (4) whether children differ in their ability to learn from S- or O-frames for reasons which may be related specifically to a marked/unmarked language learning principle.

Method

Subjects

The subjects were 56 children at a primary school in London. The chronological age range of the subjects was 5:0 to 11:6 at the time of the study. No child was included whose verbal IQ as assessed by the English Picture Vocabulary Test (EPVT) was below 60, or whose parents were not native English speakers.

Materials

The materials used in assessing the child's stage on this structure consisted of a stopwatch and two colourful, rubber hand-puppets—one the head of a duck and the other the head of a wolf. These are constructed in such a way that the thumb operates the lower jaw of the puppet, and the opposing fingers operate the upper jaw. This enables the child to control the puppet's mouth, so as to show the wolf or the duck talking, biting, etc.

In the experiment proper, in addition to these two hand-puppets, there were puppets of a monkey and of Pluto the dog. These are the same in their construction as the wolf and the duck. A wooden toy puppet theatre was used for the presentation of the puppet shows. It has a black back curtain which prevents distraction by objects behind the theatre. A red front curtain operated by pull strings on each side isolates the beginning and ending of the action of the puppet shows.

Procedure

Each child was tested individually in two separate sessions a few days apart. The median interval between the two sessions was 3 days.

In the first session, the child was given Raven's Coloured Progressive Matrices, and the English Picture Vocabulary Test (EPVT). (The relationships between scores on these tests and the child's stage on this structure will be treated, along with other developmental data, in a separate paper.) The child then undertook the adjective assessment test. In order to do this, he was first shown how to operate the two puppets, one on each hand, and he was allowed a few minutes to play with these puppets. During this warm-up the child was asked to perform the actions appropriate to two simple sentences spoken by the experimenter: " 'The wolf bites the duck'. Show me that", and " 'The duck bites the wolf'. Show me that". In the course of this warm-up, the experimenter indicated that the animal doing the biting should bite the other on the ear or top of the head so that the experimenter could clearly see which animal was doing the biting. The adjective assessment test itself consisted of 12 sentences of the structure being studied, "The wolf (or duck) is ——————— to bite", spoken by the experimenter. These were the same adjectives used by Cromer (1970). The first eight sentences used unambiguous adjectives in the space—four S-type adjectives: happy, anxious, willing and glad; and four O-type adjectives: tasty, easy, hard and fun. A counterbalanced design controlled for order effects in terms of the type of adjective first used in the structure, and for any effects due to the particular animal used with each adjective—half the subjects getting the duck as surface structure subject for a particular adjective and half getting the wolf for that adjective. These eight sentences were followed by four sentences with ambiguous (A-type) adjectives (bad, horrible, nice and nasty), similarly controlling for animal used with a particular adjective. This set of A-adjectives was included to insure that the youngest children were treating all adjectives as S-adjectives, and not just the two marked forms. The time taken for each answer was measured by stop-watch.

The second session, a few days later, consisted of two tasks. The first was a repetition of the adjective assessment test in the same order, except that the four ambiguous adjectives were omitted, in order to discourage any set to respond randomly. This second assessment was carried out as earlier studies had indicated inconsistencies from day to day which were interpreted as being part of Intermediate status as regards this structure (Cromer, 1970). This was followed by the experiment of introducing "new" nonsense words during two puppet shows. The purpose of using puppet shows as a technique for presenting the new words was to make the presentation more like the natural situations in which children are exposed to new words in the course of conversation. For these puppet shows, the child removed the wolf and duck puppets from his hands and moved to a different part of the room where he was seated in front of the puppet theatre. The experimenter explained that the child was now going to see some puppet shows:

> These puppet shows are very different because they have some funny words in them— words you don't know. And the reason you don't know them is because they're not really words at all! They're just made up for a game. They're "pretend" words, but you listen for them anyway.

The child pulled the string to open the curtain and the experimenter presented the puppet show.

The puppet shows each used only two new words. This limitation on the number of items to be learned was made in order to restrict the burden on memory. In addition, each puppet show was kept short—each lasting approximately 5 min. One puppet show was used for two S-type nonsense adjectives but in different differentiating frames, and the other used two O-type nonsense adjectives again in two different differentiating frames. A counter-balanced design of eight orders insured that the nonsense words were used about equally in all frames, and that about half the children received the S-type adjective puppet show first, and the other half the O-type first. These groups were further subdivided on the test structure, so that about half of each order was tested using the wolf as the named animal, the other half having the duck as the named animal. The learning frames came near the end of each puppet show while the child was involved in the ongoing story. These frames are given below with the words used in one of the counterbalanced design orders:

Pertinent Portion of S-type adjective Puppet Show

Pluto: Thank you so much. I'm *skelgious* that you read me the title. Yes, I'm *skelgious* that you read me the title.

Monkey: Think nothing of it. I like to read. I'm always *ispy* to read to you. Yes, I'm always *ispy* to read to you.

Pertinent Portion of O-type adjective Puppet Show

Monkey: Of course. Now come over here where I can see it. Ah, there. I always find that reading neck-watches is *riffable*. Yes, reading neck-watches is *riffable*. [Note: in the puppet show story, a neck-watch was a watch which Pluto wore around his neck and thus was unable to see in order to tell the time.]

Pluto: What time is it, then?

Monkey: (states a time).

Pluto: Thank you very much.

Monkey: My pleasure. I like telling time. I find that for me to tell you the time is *narcious*. Yes, for me to tell you the time is *narcious*.

At the end of the first puppet show, the child and experimenter returned to the table where the child again put the wolf and duck puppets on his hands. The experimenter then said, "Now show me, who does the biting in these: 'The wolf is skelgious to bite'. Show me that." And " 'The wolf is ispy to bite'. Show me that". After this, the child again removed the puppets, saw the second puppet show, and then took the test for the two nonsense adjectives used in that puppet show. The response latencies for all answers were timed by stop-watch.

Results

Classification of the children

The children were classified according to two different aspects of their answers. The first of these was their initial performance on the adjective assessment test using normal English words. If the child answered all sentences by showing the surface subject as being the actor, thus correctly interpreting those sentences using S-adjectives, but consistently misinterpreting all sentences using O-adjectives, it was said that he was using a "primitive" rule for recovering deep structure meaning. Those who gave mixed answers, sometimes showing the named animal as doing the biting and sometimes showing the "other"—some of these being wrong—were labelled Intermediate. Those children who interpreted all eight unambiguous adjectives in the adult fashion were called "Passers". The second aspect of classification was based on the child's performance on being retested on the same adjectives. Day one Primitive Rule Users who on the second day gave mixed answers, or Day one Intermediates who gave all primitive rule answers on the second day were classified as Low Intermediates. There were only three Low Intermediates and as later results showed that they performed mainly like consistent primitive rule users, they have been combined with that group. Day one Intermediates who gave all adult answers on the second day, or Day one Passers who "slipped back" into giving mixed answers, were classified as High Intermediates. The final grouping of the 56 children then was:

Primitive Rule Users (used the primitive rule for interpreting all eight adjectives on one or both days) 12

Intermediates (gave mixed answers on both days) 21

High Intermediates (gave mixed answers on one day; "Passed" all
    adjectives on the other) **9**

Passers (gave adult interpretation to all eight adjectives on both days)   **14**

Response strategies

In the two puppet shows, the child gave a total of four responses showing which animal was doing the biting when the new words were used in the test structure—two for the S-frame puppet show and two for the O-frame puppet show. If a child showed the surface structure subject to be doing the biting as his response for all four instances, he was considered as using the primitive rule. If he always answered by showing the "other" animal to be doing the biting, he was rated as using a consistent fixed strategy which can be labelled the "O-rule". The remainder of the children gave mixed answers, i.e. did not use fixed strategies, and some of these children performed correctly on all four new adjectives. The numbers and percentages of children performing in these ways are shown in Table I which is a two-way classification of the children on the nonsense words

TABLE I

Responses to "new" (nonsense) adjectives heard in a puppet show, by the four groups of children as defined by their performance on the normal adjectives

| Groups | Response types to "new" nonsense adjectives | | | | |
	Primitive rule	O-Rule	Mixed answers	All correct	Total
Primitive rule users	9 (75·0%)	1 (8·3%)	2 (16·7%)	0 (0·0%)	12 (100·0%)
Intermediates	5 (23·8%)	8 (38·1%)	8 (38·1%)	0 (0·0%)	21 (100·0%)
High Intermediates	0 (0·0%)	2 (22·2%)	7 (77·8%)	0 (0·0%)	9 (100·0%)
Passers	0 (0·0%)	5 (35·7%)	7 (50·0%)	2 (14·2%)	14 (100·0%)

$S = 534$, $\text{Tau}_c = 0\cdot454$, $Z = 4\cdot21$, $P < 0\cdot001$.

and on the real word adjective assessment test. As can be seen, most Primitive Rule Users (75·0%) continued to use the primitive rule to assign meaning to the new adjectives. Some Intermediate children (23·8%) also used the primitive rule to assign meaning to the new adjectives, but many used the O-rule (38·1%) or gave a mixed set of answers (38·1%); none performed correctly on all four new adjectives. Children who were High Intermediates no longer used the primitive rule as a strategy in answering the new adjectives; two used the O-rule, but most (77·8%) gave mixed answers, none getting all four adjectives correct. Passers similarly never used the primitive rule. Several used an O-rule strategy, but most gave mixed answers, with two children (14·2%) getting all four adjectives correct, being the only children of any group to do so. Using Kendall's Tau (Kendall, 1955), the trends in Table I were found to be significant at the 0·001 level ($S = 534$, $\text{Tau}_c = 0\cdot454$, $Z = 4\cdot21$). It should also be noted that no child

in any group used position strategies or particular animal strategies, in giving his answers.

Performance by children giving mixed answers

It has already been seen that there is a trend away from use of fixed strategies to interpret new words as one moves up the scale from Primitive Rule Users to Passers. One can also pose the question: when the responses by children using such predetermining strategies are eliminated, do the groups differ in their ability to extract structural information from the differentiating transformational frames? Only two Primitive Rule Users gave mixed answers, but enough Intermediates, High Intermediates, and Passers did so to analyse their errors. Table II shows the mean number of errors by children in each of these latter groups. As the

TABLE II

Mean number of errors on each adjective type by children in three groups not using a fixed rule strategy on nonsense words heard in a puppet show. (Errors possible on each adjective type = 2·00; Total errors possible = 4·00)

| Group | Adjective type | | Total |
	S-Adjectives	O-Adjectives	
Intermediates	1·125	1·125	2·25
High Intermediates	1·14	1·29	2·43
Passers	0·89	0·44	1·33

Fmax was not significant (Fmax = 4·04; $df = 2, 8$), a two-way analysis of variance, with the three groups as one factor and the two types of adjectives (S- and O-adjectives) as the other was carried out. There was a significant difference between the groups ($F = 3·80$; $df = 2, 21$; $P < 0·05$). Differences between adjective type were not significant ($F = 0·70$, $df = 1, 21$, N.S.) nor were interaction effects ($F = 2·29$; $df = 2, 21$, N.S.). The mean number of errors for the Intermediate and High Intermediate groups, 2·25 and 2·43 did not significantly differ from chance guessing (2·0 errors). The mean number of errors by Passers who gave mixed answers was 1·33, which while being in a direction toward more correct answers, similarly did not reach a significant difference from chance ($t = 1·79$; $df = 8$). However, this mean number of errors by Passers was significantly less than that for Intermediates ($t = 3·15$; $df = 2, 21$; $P < 0·05$), for High Intermediates ($t = 3·63$; $df = 2, 18$; $P < 0·01$), and for Intermediates and High Intermediates combined ($t = 3·98$; $df = 2, 18$; $P < 0·01$) as analysed by the Scheffé method.

If one looks more closely at the particular errors made by Passers, one finds that although interaction effects were not significant, and scores were improved (i.e. fewer errors) for both types of adjectives, the major contributing factor to the main effect was the low number of errors on O-adjectives. The mean number of errors on S-adjectives by Passers (0·89), though lower than the mean number of errors by the other groups, is nevertheless not significantly different from chance.

But the mean number of errors on O-adjectives (0·44) is significantly different from chance ($t = 2·29$; $df = 8$; $P < 0·05$, one-tail) and significantly different from the mean number of errors on O-adjectives by the Intermediate (1·125) and High Intermediate (1·29) groups at the 0·05 and 0·02 levels respectively ($t = 2·40$, $df = 15$; $t = 2·63$, $df = 14$).

Further analysis of errors revealed that there was no tendency to respond to any of the nonsense words consistently as an S- or O-adjective, no effects from viewing the S- or O-adjective puppet show first, and no effect due to which animal was named with the particular adjective. There was also no differential effect on learning due to the differentiating frame used.

Response latencies

It can be hypothesized that the children using a fixed rule strategy for interpreting new words should take a shorter time to respond than children who did not use such fixed rules. Comparing children across groups who used fixed rules with those who did not, it was found that 48·3% of rule-users had a median response latency on the four puppet show words greater than their own median response time on the preceding four normal adult words (willing, hard, glad and fun for all orders) of the adjective assessment test. However, of children who did not use such rules, 80·8% had longer median response times on the puppet show words. The difference between these percentages (based on 29 rule users and 26 non-rule children, excluding one tie) was significant at the 0·01 level, one-tail ($Z = 2·50$), by the significance of difference of Proportions Test (Blalock, 1960). Within the Primitive Rule User group, there was a slight tendency to increase in speed of response to the puppet show words. Within the Passers group, there was a tendency to react more slowly than on normal adult words, even by those children who had used the O-rule to interpret the new nonsense words.

Discussion

The results from the puppet show experiment reveal two developmental trends in the way children treat new linguistic information relevant to the "John is eager/easy to please" structure, as well as some facts relevant to the principles children use in extracting such information. The first significant trend is that fewer children in each successively higher classification level rely on particular fixed strategies in interpreting the "new", nonsense adjectives in the test structure. Furthermore, the particular strategies used by children differ at the various classification levels. Thus, Primitive Rule Users continued to use the primitive rule. A few Intermediates also used the primitive rule, but most used either an O-rule as a strategy for interpreting sentences containing the new words, or gave mixed answers. By the time the child is on the verge of adult usage (High Intermediates), he has stopped using the primitive rule; some high intermediate children used the O-rule, but most gave mixed answers. Some Passers, like the High Intermediates, also used the O-rule, but again most gave mixed answers to the new words. It was also found that children who did not use fixed rules for interpreting the new puppet show words took longer to respond to these new

words than to normal adult words, in a significant number of cases, whereas the time taken to respond by children using fixed rules did not show such differences. It is as if the children who did not use these strategies were trying to use information (structural or otherwise) about the particular new word, and these processes are quantitatively measurable.

The second trend found was that amongst the children who did not use fixed rules for interpreting the new words, those children classified as Passers on normal adult words were more successful at extracting the essential information about the new words than children of lower classification levels. Thus, when only children who had given mixed answers were compared, it was found that Passers made significantly fewer errors than did Intermediates or High Intermediates.

The improvement by Passers, however, was mainly due to significantly fewer errors on the O-adjectives. It may seem strange that with information supplied about both S- and O-types of adjectives in this structure, children with the ability to utilize that information did so more readily for the type of adjective which violates the principle by which they had located deep structure at an earlier age.

There is evidence, however, that marking both forms which indicate a contrast, when only one need be marked, proves confusing for children acquiring language. In Japanese, for example, both the direct object and the indirect object are marked by a postpositional element. (In English, identification of direct and indirect object is accomplished by word order, so that "John gave the boy the tiger" and "John gave the tiger the boy" are clearly different in terms of the direct and indirect object relationships.) McNeill, Yukawa and McNeill (1971), by using sentences in which various elements were left unmarked, have shown that Japanese children who have reached the level of being aware of the postpositional marking of objects, find the marking of the direct object confusing. They claim that it has been found to be a universal linguistic principle that when only one member of the pair "direct object"/"indirect object" is marked, the marked member is the indirect object. They found that in accordance with this universal principle of marking, the Japanese children in their experiment sought unmarked direct objects and marked indirect objects. Thus, children who had learned some but not all of the Japanese marking system were most accurate when only the indirect object was marked, doing less well when the direct object was also marked.

In a similar fashion in this experiment, the additional "marking" of S-adjectives may be confusing if children expect to find marked only those adjectives which serve as a clue that the deep structure has been transformed. Thus, providing the "marking" of O-adjectives by using them in related transformations where S-adjectives are excluded, enabled those children who had restructured their adjective categories and who were able to interpret the meaning of this linguistic structure in an adult fashion when normal words were used, to perform correctly on these new nonsense adjectives. However, when S-adjectives were similarly marked for S-interpretation, these children did less well. On the other hand, children who had given up using the primitive rule, but who had not yet properly restructured their adjective categories in the adult manner (Intermediates and High Intermediates), were unable in either case to benefit from the structural information in the puppet show sentences and, when not using a fixed strategy (such as the

O-rule) to interpret the test sentences, they merely performed at chance levels on both types of new words.

I wish to thank Mr C. Glover, headmaster, and the teachers of Holy Trinity School for their kind co-operation in providing subjects and facilities.

References

BACH, E. and HARMS, R. T. (Eds) (1968). *Universals in Linguistic Theory*. New York: Holt, Rinehart & Winston.

BELLUGI, U. (1967). The acquisition of the system of negation in children's speech. Unpublished doctoral dissertation, Harvard University.

BELLUGI-KLIMA, U. (1969). Language acquisition. Paper presented at the Wenner-Gren Foundation for Anthropological Research in the symposium on Cognitive Studies and Artificial Intelligence Research, Chicago.

BLALOCK, H. M. (1960). *Social Statistics*. New York: McGraw-Hill.

BROWN, R., CAZDEN, C. and BELLUGI-KLIMA, U. (1969). The child's grammar from I to III. In HILL, J. P. (Ed.), *Minnesota Symposia on Child Psychology*, Vol. 2. Pp. 28–73. Minneapolis: University of Minnesota Press.

CHOMSKY, C. (1969). *The Acquisition of Syntax in Children from 5 to 10*. (Research Monograph No. 57.) Cambridge, Massachusetts: M.I.T. Press.

CHOMSKY, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, Massachusetts: M.I.T. Press.

CROMER, R. F. (1970). "Children are nice to understand": Surface structure clues for the recovery of a deep structure. *British Journal of Psychology*, **61**, 397–408.

DONALDSON, M. and BALFOUR, G. (1968). Less is more: a study of language comprehension in children. *British Journal of Psychology*, **59**, 461–71.

GREENBERG, J. H. (1966a). Language universals. In SEBECK, T. A. (Ed.), *Current Trends in Linguistics*, Vol. III. Pp. 61–112. The Hague: Mouton & Co.

GREENBERG, J. H. (1966b). *Universals of Language*. Cambridge, Massachusetts: M.I.T. Press.

HOUSEHOLDER, F. W., JR., ALEXANDER, D. and MATTHEWS, P. H. (1964). Adjectives before *that*-clauses in English. (Linguistic Research Project, Indiana University.)

JAKOBSON, R. (1968). *Child Language Aphasia and Phonological Universals* (original German edition, 1941). The Hague: Mouton.

JOOS, M. (1964). *The English Verb*. Madison: The University of Wisconsin Press.

KENDALL, M. G. (1955). *Rank Correlation Methods*. New York: Hafner Publishing Company.

KESSEL, F. S. (1970). The role of syntax in children's comprehension from ages six to twelve. Monographs of the Society for Research in Child Development, **35** (6, Serial No. 139).

McNEILL, D. (1970). The development of language. In Mussen, P. H. (Ed.), *Carmichael's Manual of Child Psychology*, Vol. I. Pp. 1061–161, New York: John Wiley & Sons.

McNEILL, D., YUKAWA, R. and McNEILL, N. B. (1971). The acquisition of direct and indirect objects in Japanese. *Child Development*, **42**, 237–49.

MILLER, G. A. and McNEILL, D. (1969). Psycholinguistics. In LINDZEY, G. and ARONSON, E. (Eds). *The Handbook of Social Psychology* (2nd ed.), Vol. 3. Pp. 666–794. Reading, Massachusetts: Addison-Wesley.

REMEMBERING COMPLEX SENTENCES

ALEXANDER J. WEARING

Yale University

Recall of complex sentences at two retention intervals was examined, using sentences which varied simultaneously in three ways, being either active or passive, of low or high Yngve depth, and predictable or unpredictable. Recall of any particular sentence was cued with either the logical subject, logical object, verb or adverbial phrase noun. In general, unpredictable sentences were recollected better than predictable ones, low Yngve depth sentences were recalled better than high Yngve depth ones, and passive sentences were retained better than active ones. The most effective cue was the object, followed in turn by the subject and the adverbial phrase noun, with the verb being least effective. Implications of these findings are discussed.

Introduction

Processing a sentence into memory seems to depend on at least three factors, surface structure (Johnson, 1968; Garrett, Bever and Fodor, 1966; Ammon, 1968; Martin and Roberts, 1966), deep structure (Garrett and Fodor, 1968; Blumenthal, 1967; Savin and Perchonock, 1965; Sachs, 1967), and the semantic relations within the sentence (Clark and Clark, 1968; Fillenbaum, 1966; Rosenberg, 1966, 1968; Rubenstein and Aborn, 1958; Slamecka, 1964).

The role of surface structure in the *memory* (as opposed to perceptual segmentation) of sentences is a matter of contention. The most notable hypothesis relating surface structure to sentence memory is due to Martin and Roberts (1966) who suggest that processing difficulty is a function of the number of left branches in a sentence. Since their hypothesis was derived from the phrase structure grammar of Yngve (1960) it is usually referred to as the Yngve depth hypothesis. Evidence supporting the depth hypothesis has come from Martin and Roberts (1966), Roberts (1968) and Wang (1970), but experiments due to Perfetti (1969a,b) and Rohrman (1968) have cast some doubt on the validity of the depth hypothesis. Perfetti (1968b) found that lexical density (by which he means the ratio of content words to the total number of words in a sentence) was related to variations in sentence retention, whereas Yngve depth was not. Rohrman (1968) found that when surface structure was held constant, variations in deep structure led to differences in the recall of English nominalizations. (However, see Paivio (1971) and Wearing (1971) for some misgivings about his findings.)

A number of variables in these experiments were not systematically controlled. Neither length nor structural frame were the same for the sentences in the Martin and Roberts and the Perfetti studies. Only Perfetti controlled for the number of content words. Most studies have looked at immediate memory only, although

when one discusses memory for sentences, presumably some time delay greater than 30 sec or even 2 or 3 min is also implied. Finally, the role of semantic variables such as those suggested by Rosenberg (1966, 1968) has been ignored almost completely.

Wearing (1970) attempted to meet some of these problems by using sentences in which the deep structure, the surface structure, and semantic factors were simultaneously co-varied. After getting subjects to learn a list of 64 sentences in which each sentence varied simultaneously on three dimensions, being either active or passive, simple or complex (in an Yngve sense), and predictable or unpredictable, he presented them with a list of 128 sentences containing the original 64 sentences plus 64 new sentences, and required the subjects to identify the original and the new sentences. He found that the deep structure variable (active or passive voice) exerted no influence on recognition, whereas Yngve depth and semantic predictability did, although contrary to previous results, unpredictable sentences were retained better than predictable ones.

However, there are reasons for suspecting that the scope of these results is limited. Each new sentence was derived from one of the old sentences in that either the location of the adverbial phrase was changed or a content word was replaced with a near synonym. As a measure of retention, this procedure is open to the objection that the similarity between different old and new sentence pairs may have varied in magnitude and that such variations may have been partly responsible for the results obtained. Moreover, either a new sentence had to precede its corresponding old sentence or *vice versa*, and it is conceivable that this may have resulted in interference between recognition responses. Finally, recognition and recall may be tapping different memory traces (Adams, 1967; Adams and Bray, 1970; Wearing, 1971). Kintsch (1970) summarizes much of the evidence which suggests that recall and recognition involve different processes, and indicates that there is no necessary reason why results obtained with a recognition procedure should hold in a recall experiment.

Ideally, what is required is some more direct method of measuring sentence retention. Because of the large number of sentences originally learned (64), an appropriate procedure is one in which each sentence is cued by one of its constituent words. Not only does a cueing procedure permit a more efficient assessment of retention, but it may also provide an index of the way in which a sentence is stored since the efficacy of a cue will presumably depend on its location in the pattern of intrasentence associations (Ammon, 1968; Blumenthal, 1967; Kennedy and Wilkes, 1968, 1969; Weisberg, 1969).

In order to assess the reliability and the generality of Wearing's (1970) results, an experiment was run in which his experimental materials and general design were used, but a cued recall procedure employed to determine retention.

Method

Acquisition sentences

These consisted of the 64 experimental sentences used (and described in detail) by Wearing (1970) plus four additional buffer sentences at the beginning and end of the list. Since each sentence was either active or passive, predictable or unpredictable, or simple

or complex, there were eight different types in all (see examples below). Each sentence type was represented by eight exemplars.

Predictability was defined in terms of the real world likelihood of occurrence of the situation described by the sentence, and the characterizing of a sentence as predictable or unpredictable was validated by the judgments of an independent group of 46 subjects. The high Yngve depth (complex) sentences were formed by placing the adverbial phrase before the main verb, and low Yngve depth sentences formed by placing the adverbial phrase after the main verb. Since Yngve depth is a simple function of the number of left branches of the structural tree, no particular attempt was made to control the relationship between the adverbial clause and the rest of the sentences. The complex sentences had a mean Yngve value of 1·69, and the simple sentences had a mean Yngve value of 1·40.

Each sentence possessed eight content words and none of the sentences shared any common content words (nouns, adjectives, verbs and adverbs). Word frequency and sentence length were approximately equal across the eight sentence types. An example of each type of sentence is shown below. The bracketing indicates the structure from which the Yngve depth was computed.

Active-Predictable-Simple. ((The (brown rabbit)) ((anxiously eyed) (((the (circling eagle)) (in (the (clear sky)))))).

Active-Predictable-Complex. (((From (his (fast jeep))) (the (county sheriff))) ((quickly drew) (his (trusty pistol)))).

Active-Unpredictable-Simple. ((The (gray kangaroo)) (((joyously trampled) (the (growing barley))) (with (his (clumsy tread))))).

Active-Unpredictable-Complex. (((To (the (arrogant marquis))) (the (famous chef))) ((seriously handed) (some (burnt toast)))).

Passive-Predictable-Simple. ((Foreign policies) (((are (wrongly moulded)) (by (verbal radicals))) (in (the (Eastern towns))))).

Passive-Predictable-Complex. (((At (the (busy airport))) (the (tedious delay))) ((was (soon resented)) (by (the (alien traveller))))).

Passive-Unpredictable-Simple. ((The (ugly boss)) ((was (grandly entertained))) (by (the (poor cleaner))) (with (Swiss liqueur))).

Passive-Unpredictable-Complex. (((Into (the (shocked onlookers))) (the (fresh peach))) ((was (deftly socked)) (by (the (surly dentist))))).

Priming cues

The four possible cues for each sentence were either the logical subject (and its adjective), the verb (and its adverb), the logical object (and its adjective) or the noun (and its adjective) from the adverbial phrase. For the examples given above, the cues for the first sentence were *brown rabbit, anxiously eyed, circling eagle, clear sky,* and for the second, *country sheriff, quickly drew, trusty pistol,* and *fast jeep.*

Experimental design

All of the sentences were learned by every subject, then one group of 36 subjects was given a recall test immediately, while a second group of 36 subjects was tested after a retention interval of 48 h. Presentation order at acquisition was counterbalanced across subjects, but everyone had the same (new) recall order.

Since each sentence possessed four possible cue terms, four lists of cues were constructed so that each sentence was represented exactly once in each list, and all the cues from each sentence appeared once. Both groups of 36 subjects were divided into four subgroups of nine subjects each and at recall each subgroup was given a different list of cues, although every subject had to recall the same 64 sentences. This between subjects variable is hereafter referred to as *Cue List.*

Apparatus

The experiment was run on the PLATO equipment at the University of Illinois (Blitzer, Hicks, Johnson and Lyman, 1967). The PLATO system (as used in this experiment)

comprised 20 terminals (consisting of a typewriter keyset and CRT display) individually controlled by a computer. This capability allowed presentation in the order of the sentences to be varied.

Subjects

Seventy-two undergraduate males from the University of Illinois served as paid subjects. They received $3.00 each for participating in the experiment.

Procedure

After the subjects were seated at their terminals and instructed in general terms that they would have to learn and remember some sentences, they were presented with three practice sentences and given 31 sec to write down each sentence on a prepared form which had been placed next to their PLATO keyset. The reasons for making subjects write out every sentence were twofold. First, and most importantly, to ensure that he paid at least some attention to every word in the sentence, instead of just concentrating on major content words. Second, to prevent (to some degree) the subject from using the pattern of words on the screen (each sentence had to be written out over three rows) to help him learn.

Following the practice sentences, instructions were shown which both checked that the subjects understood the procedure, and prepared them for the main acquisition list, consisting of 64 experimental sentences preceded and followed by two buffer sentences. As in the practice trials, each sentence was shown for 31 sec.

At retention, the subjects were told that they would be presented with two cue words and that they then had to write out the sentence from which the cue words were taken. They were then shown a practice series of three cue word pairs for 35 sec each, during which time they attempted to recall and write down the appropriate sentence. As an acquisition, 3 sec before each trial ended, the warning "three seconds left" was projected at the bottom of the screen. Following the practice items, it was ascertained that the subjects understood the procedure. Then using the same procedure as for the practice items, the main recall test began during which 64 cue pairs were presented to each subject.

Results

For each sentence the number of words correctly recalled was scored. As there were two sentences in each of the eight sentences-type × four cue-type categories, the number of words correct (excluding the two cue words) in both sentences was added together. The resulting score was the number of words correct per experimental condition, and there were 32 such scores for each subject.

TABLE I

Mean number of words correctly recalled as a function of voice, predictability, Yngve depth and cue type

| | Active | | | | Passive | | | | Row |
| Cue type | Predictable | | Unpredictable | | Predictable | | Unpredictable | | mean |
	simple	complex	simple	complex	simple	complex	simple	complex	
Subject	4·64	4·58	5·89	4·21	4·53	3·97	6·01	5·25	4·88
Verb	2·71	3·12	4·69	4·07	3·58	4·44	3·99	3·57	3·77
Object	4·46	4·53	6·47	3·71	5·28	5·75	7·04	5·46	5·34
Adverb noun	3·93	4·36	6·83	3·96	3·93	3·03	5·69	4·79	4·57
Column means	3·93	4·15	5·97	3·99	4·33	4·30	5·68	4·77	

These scores were analysed by analysis of variance. Of the main effects, the decline in performance over the 48 h retention interval was only moderate from a mean of 5·16 to 4·12 ($F = 3·53$; $df = 1,64$; $P < 0·10$); unpredictable sentences were remembered better than predictable ones with a mean of 5·10 compared with 4·18 ($F = 41·52$; $df = 1·64$; $P < 0·0001$); simple sentences were remembered more accurately than complex ones with a mean of 4·98 compared with 4·30 ($F = 39·69$; $df = 1,64$; $P < 0·0001$); and passive sentences were remembered better than active ones with a mean of 4·77 as against 4·51 ($F = 4·21$; $df = 1,64$; $P < 0·05$). There was a significant difference in effectiveness between the four cues ($F = 20·18$; $df = 3,192$; $P < 0·0001$); the object was the most effective in promoting recall, followed by the subject, adverbial phrase, noun, and verb, in that order. The means of this main experimental effect are shown in Table I, which summarizes that outcome of the experiment.

Although *Cue list* itself was not significant, it did enter into two significant triple interactions with cue type and predictability ($F = 8·66$; $df = 9,192$; $P < 0·0001$), and cue type and Voice ($F = 4·31$; $df = 9,912$; $P < 0·0001$). In both cases these interactions arise because of the variation in effectiveness of the adverbial phrase noun as a cue. This variation may be due in part to the fact that the locus of the adverbial phrase depended on the Yngve depth of the sentence. It may also be the case that adverbial phrases differed with regard to their importance for understanding the propositions of their respective sentences. However, the variance accounted for by these triple interactions is small relative to that accounted for by the main effects and thus need not particularly influence the interpretation of the latter. Nevertheless, experiments (such as this one) in which the independent variables are confounded with lexical content, tend to result in higher order interactions which arise from the fortuitous combination of particular words in particular sentences.

TABLE II

Mean number of words recalled correctly as a function of voice or and cue type

Voice	Cue type				Row mean
	Subject	Verb	Object	Adv. Phrase Noun	
Active	4·82	3·65	4·79	4·77	4·51
Passive	4·49	3·90	5·88	4·36	4·77
Column mean	4·88	3·77	5·34	4·57	

Three interactions among the experimental variables deserve mention. Predictability interacted with retention interval ($F = 8·90$; $df = 1,64$; $P < 0·005$); Voice (active or passive) interacted with cue type ($F = 5·33$; $df = 3,192$; $P < 0·005$) (see Table II); and Yngve depth interacted with predictability ($F = 22·34$; $df = 1,64$; $P < 0·0001$) (see Table III). The interaction of predictability with retention interval arises because unpredictable sentences were remembered much better at immediate recall (a mean of 5·84 compared with 4·49) than 48 later (a mean of 4·36 compared with 3·87), although the difference remains significant.

TABLE III

Mean number of words recalled correctly as a function of Yngve depth and predictability

| | Yngve Depth | | Row |
	Simple	Complex	mean
Predictable	4·13	4·22	4·18
Unpredictable	5·82	4·38	5·10
Column mean	4·98	4·30	

The interaction of Voice with cue type stems from the significantly greater efficacy of the object as a cue in passive as opposed to active sentences. The interaction of predictability with Yngve depth is due to the relatively better recall of simple unpredictable sentences.

Discussion

These results, using quite a different method of assessing retention from that of Wearing (1970), generally confirm his findings, with unpredictable sentences being remembered better than predictable ones and complex sentences being recalled better than simple ones. Wearing (1970) found no difference due to Voice, but in the present experiment passive sentences were recalled better than active ones. However, this result (which is a weak one) is entirely due to the strong cueing effect of the object. Since one of the presumed functions of passivization is to emphasize the logical object, then one can account of its superiority as a cue in terms of the importance given to it by the passive construction (Johnson-Laird, 1968).

The greater difficulty in recalling complex as opposed to simple sentences confirms Martin and Roberts (1966) and Wearing (1970), but apparently contradicts the results of Perfetti (1969a, b). Since the sentences in the present experiment were long, they could have been relatively difficult to process, imposing a critical load on STM. The difference between simple and complex sentences may reside in the possibility that a provisional structure can be assigned sooner and more readily to low Yngve depth than high Yngve depth sentences, where the assignment of structure to an item allows it to be removed (in some sense) from working STM (Wearing and Crowder, 1971). In this experiment the difficulty would be provided by the early arrival of an adverbial phrase, which would then have to be held in STM until it could be determined which of the later arriving terms it modified. If this suggestion is correct, then the discrepancy between these results and those of Perfetti may be due to the fact that Yngve depth is not perfectly correlated with the ease of assignment of structure, a possibility demonstrated by Martin (1970). Wang (1970) using sentences up to 29 words in length, found that Yngve depth was significantly related to comprehensibility and the interaction between Predictability and Yngve depth in the present experiment supports the possibility that the load on STM is a critical factor. It is possible that unpredictable sentences place a rather heavier processing load on STM (see below) even though they are easier to remember, and it was indeed found that the

combination of complexity and unpredictability depressed the level of learning to a significant degree.

Apparently contradictory results have recently been presented by Perfetti and Goodman (1971). They varied Yngve depth by changing the location of a subordinate clause containing an adverbial noun phrase, and found that it had little predictive value in sentence recall. However, a possibly crucial factor is that their sentences were longer than the ones used here by a factor of a half, and their subordinate clauses were also much longer (five or six words as opposed to usually three). The serial position curves which Perfetti and Goodman (1971, Fig. 3) report suggest that, possibly because of their length, these phrases are processed more as strings of unrelated items, than as words in a sentence (cf. Mandler and Mandler, 1964; Wearing and Crowder, 1971).

A second finding which also apparently contradicts the present results is due to Wright (1969). She found (Experiment II) that there were no differences between high and low depth sentences when the location of the subordinate clause was moved from before the main verb to after it. Since Wright's sentences possessed fewer content words than the present ones, it is possible that the disagreement between the results is due to differences in processing load. Additionally, Wright did not assess sentence recall directly, but used the Archimedean technique developed by Savin and Perchonock (1965), thus making her findings not strictly comparable with those of the present study.

However, her results, taken in conjunction with those of Perfetti and present ones, do suggest that simply counting left branches is not an adequate index of processing load. If it were, then it would not matter what construction was responsible for a particular branching pattern, the processing difficulty would depend solely on the number of left branches. That this is not the case is attested to by the disagreements that have been referred to above, and one must conclude by saying that the notion of Yngve depth needs further explication, with perhaps rather more attention to the actual locus of processing load (Wearing and Crowder, 1971) as well as to the semantic (and structural) relationships between the various constituents.

The finding that predictable sentences were remembered more poorly than unpredictable ones runs counter to other results (Rosenberg, 1966, 1968). However, the contradiction may be only apparent.

Rosenberg devised his less predictable sentences by using words that were only weakly associated with one another, thus conjoining semantically unrelated terms such as *farmer* and *road*. But the unpredictable sentences in the present experiment used words which are negatively related like *uncouth* and *bishop*, or propositions that contain negatively related expressions such as *The wealthy earl contently lived his last years in the Negro ghetto.*

It is possible to conceive of the relationships between the terms in a sentence possessing the characteristics of both strength and sign (positive or negative). If this is the case, then one might expect that sentences containing negative relationships would be more difficult to process initially, since the listener's (or reader's) expectations will be disconfirmed, and he will have to begin processing unexpected input, but that such sentences will be easier to learn, as no new associations need

be elaborated, but only the sign of the already existing associations reversed. For unrelated expressions, on the other hand, the various semantic connections will have to be generated de novo.

Although this hypothesis accounts for the difference between the present results and those of Rosenberg, it does not necessarily predict any difference between predictable and unpredictable sentences in the present experiment. However, it is quite possible that the strength (without regard to sign) of the relationships in the present case was greater for unpredictable than predictable sentences. In addition, variables like vividness (Paivio, 1969) were not controlled, and may have been confounded. Certainly it is possible that the imagery generated by the unpredictable sentences was more vivid (Begg and Paivio, 1969). Nonetheless, whatever the precise nature of predictability, it was an effective variable, and this result indicates that experiments which investigate structure must also contain appropriate controls for semantic effects.

The differential effectiveness of the various types of cues in promoting recall has a striking parallel with Wearing's (1970) results. Using a recognition test, he found that changes in the object were most readily detected, followed in turn by changes in the subject, adverb and verb. In the present experiment, the object was the most effective cue, followed by the subject, the adverb, with the verb being least effective. The fact that a term which is well recognized is also a good cue (and vice versa) suggests that any degradation in memory occurs to the term itself, rather than to its associative connections, since otherwise one might expect that even when the term was recognized, it still might not be a useful cue. These results, of course, do not exclude the possibility that both the term and its connections are simultaneously degraded. These results confirm the findings of Martin, Roberts and Collins (1968), that word classes are differentially processed into memory. However, the poor recall of the verb is a somewhat unexpected finding. One explanation may be that verbs are less concrete than nouns, and since concrete words are generally recalled better than abstract ones (Paivio, 1969), verbs are not recalled as well. Most of the nouns used in this experiment were concrete. A second explanation may be that the verbs shared more features in common than the nouns, and possessed fewer unique features, and so were more confusable. A third (and perhaps far-fetched) possibility is that when a sentence is processed into memory, it is stored in some form which preserves the nouns as separate entities but which breaks up the verb into components which are then linked to the nouns. This last notion invokes the distinction made by Bregman and Strasberg (1968) between the semantic message and the transmission code. The present speculation is that what may be stored is the subject and object of the sentence, with the other terms in the sentence (adjectives, verb, etc.) being stored as abstract attributes of either or both of them, or, in the case of verbs, as an abstract relationship between the nouns. Consequently, although the meaning of the verb may be preserved as well as that of the nouns, its exact form may not be.

throughout the course of this study, to Rosemary Wearing for the preparation of sentences, and to Marcia Harms, Linda Kalmanek and Rosemary Wearing for assistance in the collection and the analysis of data.

References

ADAMS, J. A. (1967). *Human Memory.* New York: McGraw-Hill.

ADAMS, J. A. and BRAY, N. W. (1970). A closed-loop theory of paired-associate verbal learning. *Psychological Review*, **77**, 385–404.

AMMON, P. (1968). The perception of grammatical relations in sentences: A methodological exploration. *Journal of Verbal Learning and Verbal Behavior*, **7**, 869–75.

BEGG, I. and PAIVIO, A. (1969). Concreteness and imagery in sentence meaning. *Journal of Verbal Learning and Verbal Behavior*, **8**, 821–7.

BITZER, D. L., HICKS, B., JOHNSON, R. and LYMAN, E. R. (1967). The PLATO system: Current research and development. *IEEE Transactions on Human Factors in Electronics*, **8**, 64–70.

BLUMENTHAL, A. L. (1967). Promoted recall of sentences. *Journal of Verbal Learning and Verbal Behavior*, **6**, 203–6.

BREGMAN, A. S. and STRASBERG, R. (1968). Memory for the syntactic form of sentences. *Journal of Verbal Learning and Verbal Behavior*, **7**, 396–403.

CLARK, H. H. and CLARK, E. V. (1968). Semantic distinctions and memory for complex sentences. *Quarterly Journal of Experimental Psychology*, **20**, 129–38.

COFER, C. N. (1941). A comparison of logical and verbatim learning of prose passages of different lengths. *American Journal of Psychology*, **54**, 1–20.

FILLENBAUM, S. (1966). Memory for gist: some relevant variables. *Language and Speech*, **9**, 217–27.

GARRETT, M., BEVER, T. G. and FODOR, J. A. (1966). The active use of grammar in speech perception. *Perception and Psychophysics*, **1**, 30–2.

GARRETT, M. and FODOR, J. A. (1968). Psychological theories and linguistic construction. In DIXON, T. R. and HORTON, D. L. (Eds), *Verbal Behavior and General Behavior Theory.* Pp. 451–77. Englewood Cliffs, N.J.: Prentice-Hall.

JOHNSON, N. F. (1968). Sequential verbal behavior. In DIXON, T. R. and HORTON, D. L. (Eds), *Verbal Behavior and General Behavior Theory*, pp. 421–50. Englewood Cliffs, N.J.: Prentice-Hall.

JOHNSON-LAIRD, P. N. (1968). The interpretation of the passive voice. *Quarterly Journal of Experimental Psychology*, **20**, 69–73.

KENNEDY, R. A. and WILKES, A. L. (1968). Response times at different positions within a sentence. *Quarterly Journal of Experimental Psychology*, **20**, 390–4.

KENNEDY, R. A. and WILKES, A. L. (1969). Analysis of storage and retrieval processes in memorizing simple sentences. *Journal of Experimental Psychology*, **80**, 396–8.

KINTSCH, W. (1970). *Learning, Memory and Conceptual Processes.* New York: Wiley.

MANDLER, G. and MANDLER, J. M. (1964). Serial position effects in sentences. *Journal of Verbal Learning and Verbal Behavior*, **3**, 195–202.

MARTIN, E. (1970). Toward an analysis of subjective phrase structure. *Psychological Bulletin*, **74**, 153–66.

MARTIN, E. and ROBERTS, K. H. (1966). Grammatical factors in sentence retention. *Journal of Verbal Learning and Verbal Behavior*, **5**, 211–8.

MARTIN, E. and ROBERTS, K. H. (1967). Sentence length and sentence retention is the free learning situation. *Psychonomic Science*, **8**, 535.

MARTIN, E., ROBERTS, K. H. and COLLINS, A. M. (1968). Short-term memory for sentences. *Journal of Verbal Learning and Verbal Behavior*, **7**, 560–6.

PAIVIO, A. (1969). Mental imagery in associative learning and memory. *Psychological Review*, **76**, 241–63.

PAIVIO, A. (1971). Imagery and deep structure in the recall of English nominalizations. *Journal of Verbal Learning and Verbal Behavior*, **10**, 1–12.

PERFETTI, C. A. (1969). Sentence retention and the depth hypothesis. *Journal of Verbal Learning and Verbal Behavior*, **8**, 101–4.

PERFETTI, C. A. (1969). Lexical density and phrase structure depth as variables in sentence retention. *Journal of Verbal Learning and Verbal Behavior*, **8**, 719–24.

PERFETTI, C. A. and GOODMAN, D. (1971). Memory for sentences and noun phrases of extreme depth. *Quarterly Journal of Experimental Psychology*, **23**, 22–33.

ROBERTS, K. H. (1968). The interaction of normative associations and grammatical factors in sentence retention. Paper read at Midwestern Psychological Association meetings, Chicago.

ROHRMAN, N. L. (1968). The role of syntactic structure in the recall of English nominalizations. *Journal of Verbal Learning and Verbal Behavior*, **7**, 904–12.

ROSENBERG, S. (1966). Recall of sentences as a function of syntactic and associative habit. *Journal of Verbal Learning and Verbal Behavior*, **5**, 392–6.

ROSENBERG, S. (1968). Association and phrase structure in sentence recall. *Journal of Verbal Learning and Verbal Behavior*, **7**, 1077–81.

RUBENSTEIN, H. and ABORN, M. (1958). Learning, prediction and readability. *Journal of Applied Psychology*, **42**, 28–32.

SACHS, J. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and Psychophysics*, **2**, 437–42.

SAVIN, H. B. and PERCHONOCK, E. (1965). Grammatical structure and the immediate recall of English sentences. *Journal of Verbal Learning and Verbal Behavior*, **4**, 348–53.

SLAMECKA, N. J. (1964). Acquisition and retention of connected discourse. *Journal of Experimental Psychology*, **68**, 330–3.

WANG, M. D. (1970). The role of syntactic complexity as a determiner of comprehensibility. *Journal of Verbal Learning and Verbal Behavior*, **9**, 398–404.

WEARING, A. J. (1970). The storage of complex sentences. *Journal of Verbal Learning and Verbal Behavior*, **9**, 21–9.

WEARING, A. J. (1971). Vividness and the recall of English nominalizations. *Psychonomic Science*, **22**, 121–2.

WEARING, A. J. (1971). On the Adams-Bray retrieval model. *Journal of Experimental Psychology*, **89**, 96–101.

WEARING, A. J. and CROWDER, R. G. (1971). Dividing attention to study sentence acquisition. *Journal of Verbal Learning and Verbal Behavior*, **10**, 254–61.

WEISBERG, R. W. (1969). Sentence processing assessed through intrasentence word associations. *Journal of Experimental Psychology*, **82**, 332–8.

WRIGHT, P. (1969). Two studies of the Depth Hypothesis. *British Journal of Psychology*, **60**, 63–9.

YNGVE, V. H. (1960). A model and a hypothesis for language structure. *Proceedings of the American Philosophical Society*, **104**, 444–66.

WHEN NEGATION IS EASIER THAN AFFIRMATION

P. N. JOHNSON-LAIRD AND J. M. TRIDGELL

Department of Psychology, University College, London

An experiment is reported which establishes that affirmative sentences are not always easier to grasp than negative sentences. The subjects had to make inferences from pairs of premises such as: "Either John is intelligent or he is rich. John is not rich". The task was reliably easier when the second premise was explicitly negative (as in the example) than when it was an affirmative ("John is poor"). It was most difficult when the negative occurred in the disjunctive premise and was denied by an affirmative (e.g. "John is intelligent or he is not rich. John is rich"). It is argued that it is simpler to establish that two statements are mutually inconsistent when one is the explicit negation of the other, but that the natural function of the negative is to deny.

Introduction

Negative sentences are generally harder to understand and to evaluate than affirmative sentences. This was originally demonstrated experimentally by Wason (1959) and has subsequently been confirmed by numerous investigators. Yet, in daily life, negatives hardly ever seem to be difficult. This paradox is perhaps best resolved in terms of Wason's thesis (1965) that negatives are normally used to deny plausible misconceptions. For example, the misconception, or preconception as we prefer to call it, in the case of, "John doesn't like Mary", would be that John does like Mary. It is accordingly feasible that the comprehension of such a negative in its everyday context is facilitated by the prior grasp of its preconception. Precisely this advantage is lost by the experimental presentation of the sentence in contextual isolation. There is some evidence to support this conjecture. Both Wason (1965) and Johnson-Laird (1967) were able to reduce the difficulty of negatives by ensuring that they made plausible denials. But are there any circumstances in which a negative would actually be easier than an affirmative?

Consider the logical problem of what can be inferred from the following premises:—(1) Either John is intelligent or he is rich. (2) John is not rich. It is a simple matter to appreciate that the second premise is a categorical denial of one of the alternatives in the disjunctive premise, and hence that the other alternative must be true, i.e. John is intelligent. But suppose that the negative premise is replaced by an affirmative one with the same function:—(1) Either John is intelligent or he is rich. (2) John is poor. It now seems that an extra step is required since it is necessary to appreciate that *poor* implies *not rich*. Hence, this problem is likely to be more difficult than the first one, because it contains an implicit rather than an explicit denial.

The problem is likely to be still harder if the negative is moved into the disjunctive premise:—(1) Either John is intelligent or he is not rich. (2) John is rich. It is still easy to see that there is a conflict between the categorical premise and one alternative of the disjunctive premise, but it seems very much harder to grasp what this implies. In fact, of course, one alternative is again false so the other must be true, i.e. John is intelligent.

The present study investigated all three sorts of problem: the first in which a negative is used appropriately to make a denial, the second in which an affirmative is used to make a denial, and the third in which the negative is inappropriately denied. It was predicted that the "appropriate negative" problem would be easier than the "affirmative" problem which, in turn, would be easier than the "inappropriate negative" problem.

Method

Design and materials

The subjects acted as their own controls and attempted to solve two examples of each of the three sorts of problem. The order of presentation was counterbalanced so that each of the six possible different orders for three problems, followed by its mirror image, occurred with an equal number of subjects.

In constructing the problems three boys' names and three girls' names were used, and three pairs of traits: intelligent or rich (poor), generous or beautiful (ugly), athletic or short (tall). The terms in parentheses were used, where necessary, to deny their antonyms in the first three problems encountered by subjects, they were denied by their antonyms, where necessary, in the second three problems encountered by subjects. The order of the resulting six different lexical contents was held constant over the subjects.

Subjects

Twenty-four undergraduates at University College, London were individually tested. They had no previous experience with tasks of this sort or with formal logic.

Procedure

The subjects were told that their task involved reasoning but it was not an intelligence test. They would be given a series of problems, each consisting of two premises, and they would have to determine what followed from them in virtue of logic alone. They were to make their responses as quickly as was compatible with drawing the correct conclusion.

The experimenter read aloud each problem, and timed the subjects by stopwatch from the moment that he finished reading until they uttered a response. The response was neither commented upon nor corrected. There was a single practice problem of a different logical variety to familiarize subjects with the general procedure.

Results

The mean response times for the three sorts of problem on their first and second presentations are given in Table I. The evident trend in favour of the prediction was highly reliable. Nine subjects conformed precisely to the required rank order, seven subjects partially conformed to it except that for them the "affirmative" problem was the most time consuming, and four subjects partially conformed to it except that for them the "affirmative" problem was the least time consuming. Only the results of the four remaining subjects were sufficiently far from the prediction to count against it (in terms of Kendall's P). Hence, the trend was

highly significant ($P = $ 0·001, sign test, one tail). It will be noted that the appropriate negative was less time-consuming than the affirmative for 17 out of the 24 subjects ($P < $ 0·04, sign test, one tail).

An analysis of variance was also carried out on the untransformed response times. It confirmed the significant difference between the problems ($F2,46 = $ 32·7; $P < $ 0·001), but failed to reveal any significant effects involving the lexical material. The apparent learning effect from the first to the second presentations of the problems was not significant, presumably because the "inappropriate negative" problem took more time (but yielded fewer errors) on its second presentation.

TABLE I

The mean response time (sec) for the three types of problems on their first and second presentations

| | Type of problem | | |
	Appropriate negative	Affirmative	Inappropriate negative
First presentation	4·8	6·7	8·0
Second presentation	4·2	5·5	8·8
Overall mean	4·5	6·1	8·4

A greater number of errors were made by the subjects than had been anticipated: nine errors with the appropriate negative, 12 errors with the affirmative, and 21 errors with the inappropriate negative. (Their overall mean latency was about 1·0 sec longer than that of the correct responses.) The trend is again in the predicted direction, and, since about a third of the responses were erroneous, it was considered that some statistical treatment of them was desirable. They were therefore scored according to the following conservative principles: when only one error was made by a subject, it was counted in favour of the prediction if it occurred with an inappropriate negative, against the prediction if it occurred with the appropriate negative, and neutral with respect to the prediction if it occurred with the affirmative. When more than one error was made by a subject, exactly the same scoring procedure was followed for each of them, and the overall total computed. It transpired that of the 17 subjects who committed errors, nine had positive scores in favour of the prediction, and the remaining seven subjects had neutral scores of zero. Hence, there was a reliable trend in favour of the prediction ($P = $ 0·003, sign test, one tail). The main error consisted in stating the negation of the correct conclusion.

Most of the remarks made by the subjects were symptomatic of the difficulty of the inappropriate negative. They complained that it was somehow ungrammatical or invalid to assert: "Either John is intelligent or he is not rich". They complained, more irrelevantly, that it was unclear whether the two alternatives were mutually exclusive.

Discussion

The pattern of results makes a striking contrast with the other findings on negative sentences reported in the literature. In interpretative tasks, such as matching statements to pictures, affirmatives are easier to understand than negatives (e.g. McMahon, 1963), whereas we found that in denying a statement negatives are easier than affirmatives. The reason for this contrast obviously lies in the difference between the two tasks. Pictures are likely to be encoded in a primarily affirmative fashion, and, in evaluating descriptions of them, it is natural that the aim should be to set up a one-to-one correspondence between the description and encoding. Indeed, this is a basic assumption of two independently formulated information-processing models of the task (Clark, 1971; Trabasso, Rollins and Shaughnessy, 1971). However, to grasp that one statement denies another the aim should be to establish *not* a one-to-one correspondence between them but a mutual inconsistency. This will be easiest when the two statements contradict one another, especially if one is the explicit negation of the other. It will be hardest when the two statements are merely contrary to one another, especially if they contain affirmative but autonymous predicates.

There are, of course, further complications. Within some pairs of antonyms one item can be used in a quite neutral sense (e.g. *tall*, when one asks how tall someone is), whereas the other item can be used only in a contrastive sense (e.g. *short*). This asymmetry has prompted Clark (1971) to argue that the contrastive items are implicit negatives: their meaning is defined essentially by negating their antonyms. Hence, it is plausible that they would make more natural denials of their antonyms than their antonyms would of them. A careful examination of our data failed to reveal any such difference or, indeed, any difference between the different sorts of antonyms. However, the experiment was not specifically designed to examine these factors, and it is intended to put them to a more stringent test in a further investigation.

The more crucial complication concerns the order of statements. It is difficult to see why this factor should affect the detection of a mutual inconsistency. Indeed, Greene (1969) found that it had no effect upon a task in which subjects had merely to judge whether two statements, one affirmative and the other negative, were synonymous or not. Yet our findings show that it is easy to grasp that a negative denies an affirmative, but exceedingly difficult to grasp that an affirmative denies a negative. The simplest explanation would seem to be that the subjects attempt to keep track of the attributes which apply to the given individual. Hence, with an inference from the premises—(1) Either John is intelligent or not rich; (2) John is rich—there may be a tendency to argue that the second premise negates an alternative in the first premise. Hence, John is *not* not rich; it follows that he is rich. But this, of course, is precisely the premise from which the argument started. The whole of this "double negative" inference may then start again, and continue in an almost hypnotic fashion until the subject breaks the circle by concluding that a negative follows from the premises. In the easier inferences, however, it is a simple matter to keep track of the attributes which apply to the given individual because the double negative does not occur, and thus the vicious circle does not arise.

It would be easy to suppose that the greater difficulty of negative sentences over their correlated affirmatives was one of the constants of psycholinguistics. We now know that this is not so. Perhaps it should not surprise us that the proper function of affirmatives is to make assertions, and of negatives to make denials.

References

CLARK, H. H. (1971). Semantics and comprehension. In SEBEOK, T. A. (Ed.), *Current Trends in Linguistics. Vol. 12: Linguistics and Adjacent Arts and Sciences.* The Hague: Mouton (in press).

GREENE, J. (1969). The semantic function of negatives and passives. *British Journal of Psychology*, 61, 17–22.

JOHNSON-LAIRD, P. N. (1967). One pragmatic factor governing the use of the English language. University of London: Unpublished Ph.D. thesis.

McMAHON, L. E. (1963). Grammatical analysis as part of understanding a sentence. University of Harvard: Unpublished Ph.D. thesis.

TRABASSO, T., ROLLINS, H. and SHAUGHNESSY, E. (1971). Storage and verification stages in processing concepts. *Cognitive Psychology*, 2, 239–89.

WASON, P. C. (1959). The processing of positive and negative information. *Quarterly Journal of Experimental Psychology*, 11, 92–107.

WASON, P. C. (1965). The contexts of plausible denial. *Journal of Verbal Learning and Verbal Behavior*, 4, 7–11.

COMPARISONS OF TWO MEASURES OF FREE-OPERANT AVOIDANCE UNDER TWO CONDITIONS OF RESPONSE FEEDBACK

HARRY M. B. HURWITZ,† PETER HARZEM‡ AND BEVERLEY KULIG

University of Tennessee, U.S.A.

Lever-pressing responses of 6 rats were studied under a free-operant avoidance procedure, in 6 two-hour sessions. Three of the subjects were given an added visual feedback for each response; they avoided shocks more effectively than the subjects without the feedback. Two different measures of overall performance in free-operant avoidance were applied to the results of the experiment. There was little difference between the measures: both reflected the subjects' performance equally well. Relative merits of the measures and the criteria for selecting one of them were discussed.

Introduction

This paper is primarily concerned with the problem of the measurement of performance in avoidance situations. In addition, an experiment is reported in which the effect on avoidance behaviour, of the provision of visual feedback for each response was studied. The measures to be discussed were applied to the results of the experiment. This provided an empirical basis for the comparison of the two measures.

The Problem of Measurement

In 1953 Sidman described the experimental design which is termed free-operant avoidance (or Sidman avoidance). The procedure involved a set of temporal contingencies between responses and shocks. The shocks were programmed to occur at regular intervals, the S–S intervals; each response postponed shock by a fixed period of time, the R–S interval. No discriminative stimuli were presented.

Since then the behaviour which develops under these conditions has been extensively studied. Most of the research, however, has been troubled by a problem of method; namely, the problem of finding a common basis for collating and comparing data obtained from different subjects, or from the same subject under different conditions. Typically, the results show considerable inter-subject, as well as intra-subject, variation (Sidman, 1966).

Moreover, taken singly, the commonly used measures such as the rate of response, the number of shocks received in a session, and distribution of inter-response times, fail to reflect important aspects of the subjects' behaviour. For

---

† Now at the University of Guelph, Ontario, Canada.
‡ Now at the University College of North Wales, Bangor.

example, the rate of response may remain constant while the number of shocks may increase or decrease or, conversely, the number of shocks may be stable while variations are observed in the rate of response.

In 1968 two different measures were independently published, both attempting to provide a summary of the subjects' overall performance in relation to the contingencies imposed upon that performance.

One of these, by McIntire, Davis, Cohen and Franch (1968) is called the *Index of Efficiency* (E). It is calculated from the following formula:

$$E = \frac{\text{Number of shocks}}{\text{Number of possible shocks}} \times \frac{\text{Number of responses} + \text{Minimum number of possible responses for total avoidance}}{\text{Minimum number of possible responses for total avoidance}}$$

This index takes into account (i) the number of responses in relation to the minimum number of responses which would have avoided all shocks, and (ii) the number of shocks in relation to the maximum number of shocks which could have occurred.

The other measure, by Hurwitz and Bounds (1968), is the *Proficiency of Avoidance*:

$$\text{AVPRO} = 100\left(1 - \frac{S_o}{S_{max}}\right) T_i,$$

$$\text{where } S_{max} = \frac{T_a}{R-S}, \; T_i = \frac{T_a}{T_{session}},$$

and where $T_{session}$ is the total session time, T_a is the total time that R–S intervals were in operation, R–S is the duration by which a response postpones shock, S_o is the number of shocks per session resulting from the elapse of the R–S periods.

The measure focuses on (i) the total amount of time spent in the R–S interval in proportion to the total duration of the experimental session, and (ii) the number of shocks which occurred at the end of R–S intervals in proportion to the maximum possible number of R–S shocks which could be delivered.

The Index of Efficiency has the advantage of relative simplicity, and uses data such as the number of shocks and of responses which are almost invariably recorded. The Proficiency of Avoidance requires the separate recording of two kinds of experimental events which are usually ignored, viz. (i) number of shocks which follow responses (R–S shocks), and (ii) total time spent in R–S intervals. On the other hand only one other item of information, the duration of the experimental session, is needed to calculate AVPRO. Thus the measure does not directly take into account the rate of responding. This is a major advantage since in free-operant avoidance short "bursts" of responding tend to occur, usually immediately after the delivery of shocks. Although such bursts enhance the overall rate of responding they contribute little to the time in which the subject is free of shocks.

The two measures described above were applied to the data obtained in the following experiment.

Experiment I

Bolles and Popp (1964) and Bolles and Grossen (1969) have reported that when a lever-response is immediately followed by a stimulus change, the acquisition of avoidance behaviour is facilitated. In both of these studies, however, the results are given merely in terms of the number of subjects which "learned" to respond, but no further data have been provided. In the latter study, for example, it is stated that "the three nonfeedback subjects who did learn the response were not inferior to the feedback subjects in terms of how soon they started responding at a high rate, the total number of responses made, or in terms of number of shocks received" (p. 96). But it is not shown how soon the subjects did start responding and what was the total number of responses made or shocks received. The present experiment was performed to study in greater detail the behavioural effect of providing an added visual stimulus for each response, in a free-operant avoidance situation.

Method

Subjects and apparatus

The subjects were 6 female hooded rats supplied by Spruce Farms, N.J. They were approximately 120 days old at the beginning of experimentation. Two standard operant conditioning boxes were used. In each box the lever was mounted 63 mm above floor level and responded to approximately 10 g pressure. A small houselight was mounted on the ceiling of the box. Shock was supplied from a constant current (0.8 mA, 180 VDC) power source and was delivered to the grid, floor and walls of a box via a scrambling device. Duration of shock was 0.3 sec. In one of the boxes a lever response resulted in a brief, 0.5 sec, illumination of the chamber by two pilot lights mounted on the ceiling. The boxes were housed in sound insulated chambers.

Procedures

For rats S_{1A}, S_{2A}, S_{3A} (Group A) each response on the lever initiated a 20 sec time period, the R–S interval, which if allowed to expire resulted in a brief shock. In addition, every response also produced the feedback stimulus, a 0.5 sec light flash. For rats S_{1B}, S_{2B}, S_{3B} (Group B) the lever response did not produce the feedback stimulus. In the absence of responding, shock occurred every 5 sec, the S–S interval. Each animal was given 6 two-hour sessions.

Results and Discussion

Figures 1 and 2 show the results obtained from each subject in Groups A and B respectively. In both figures the mean number of responses per min (middle row) and the mean number of shocks received per min (bottom row) are shown for all six experimental sessions. The two measures of performance, namely the Efficiency of Avoidance (McIntire *et al.*, 1968) and the Avoidance Proficiency (Hurwitz and Bounds, 1968) are given in the top row of each Figure. In order to compare the measures conveniently the Efficiency of Avoidance has been transformed according to the formula 100 (1–E). Thus in both measures higher numbers indicate greater "efficiency" of performance. For example, AVPRO is 100 when no shocks are delivered during the session.

The results showed that Group A received fewer shocks than Group B. Subjects S_{1A}, S_{2A}, S_{3A} in Group A reduced the median number of shocks over the six training sessions from 5.85 to 1; from 2.8 to 0.9; and from 3.3 to 1.2 shocks per min

respectively. Subjects S_{1B}, S_{2B}, S_{3B} in Group B reduced the number of shocks per min relatively little, that is, from 3·7 to 3·3; from 3·6 to 2·6; and from 3·7 to 2·1 respectively. The rates of response showed less consistent changes although Group A doubled their rates of response over the first five sessions. The results for Group B were less clear. Two of the subjects showed an increase over sessions, whereas one subject continued to respond at the same rate throughout all six sessions. In the final session the mean response rate for Group A was 10·3 whereas the mean rate for Group B was 6·1.

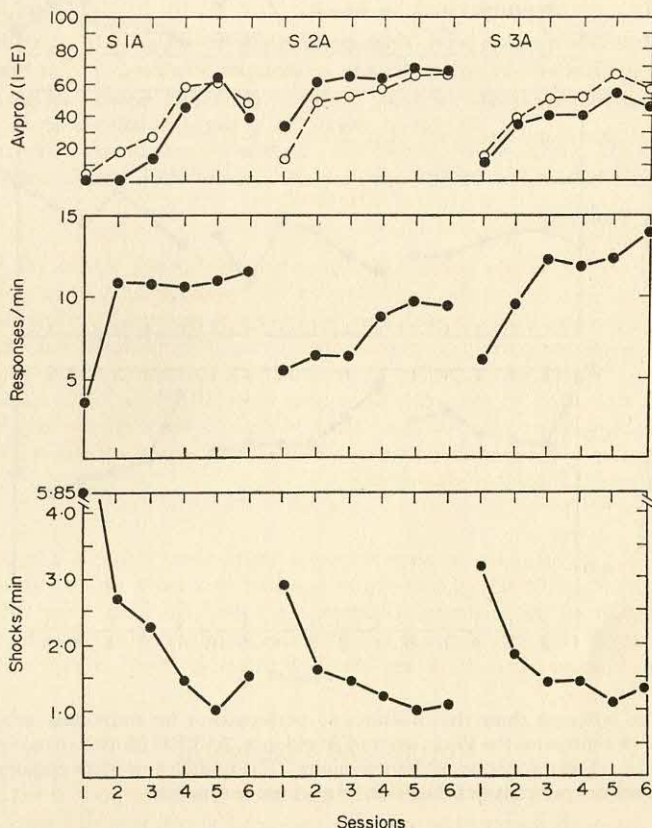FIGURE 1. The columns show the measures of performance for individual subjects, S_{1A}, S_{2A}, S_{3A}. The upper row represents the Proficiency of Avoidance, AVPRO (dotted lines) and the efficiency index, 1-E (solid lines), as a function of daily sessions. The middle row gives responses per min and the bottom row, shocks per min as a function of successive sessions.

The subjects in Group A showed a substantial increase in both the Efficiency Index and the Avoidance Proficiency as a function of training sessions. The changes over the same number of sessions for subjects in Group B were slight. In the final session the subjects in Group A achieved a mean score on AVPRO of 55 whereas Group B attained a score of 10. (A score of 100 on both measures indicates that all shocks were avoided and a score of 0 that the subjects received all shocks available during either the S–S or the R–S times.)

The performance of the two groups differed significantly ($P < 0.01$) when scores in the Efficiency Index, as well as in the AVPRO, were compared by the analysis of variance and Duncan's new multiple range test.
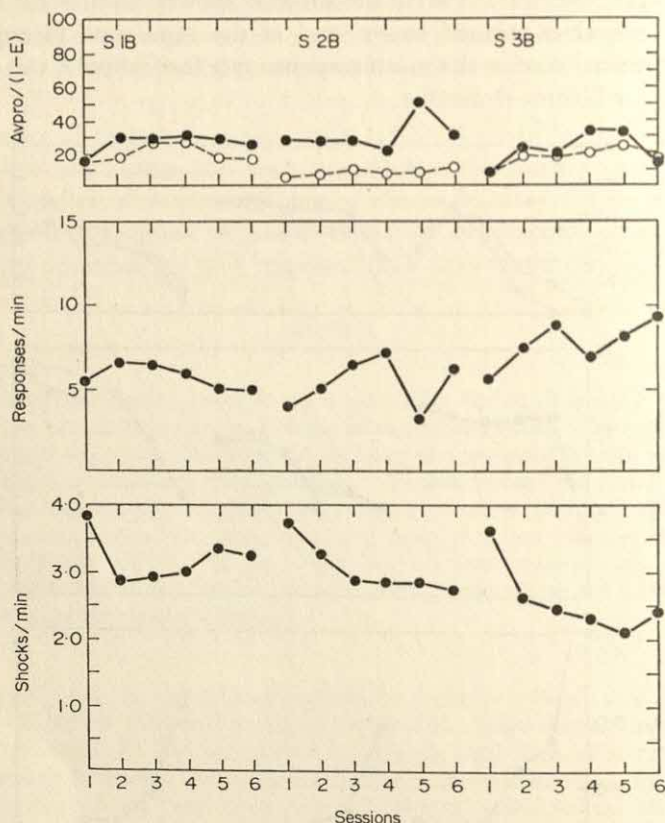

FIGURE 2.   The columns show the measures of performance for individual subjects, S_{1B}, S_{2B}, S_{3B}. The upper row represents the Proficiency of Avoidance, AVPRO (dotted lines) and the efficiency index, 1-E (solid lines), as a function of daily sessions. The middle row gives responses per min and the bottom row, shocks per min as a function of successive sessions.

There was little difference in the way the measures reflected the overall performance of the subjects. This suggests that the choice of which measure to use might, on the whole, depend on considerations such as the kinds of data collected in an experiment. There is, however, a further important point. In those studies where frequent response-bursts are observed AVPRO would be more appropriate since, as discussed above, this measure is only minimally affected by such bursts. The present study indicates that both measures are potentially useful research tools. Neither, of course, is intended to replace the more direct measures such as the rate of response, but rather to supplement them.

References

BOLLES, R. C. and POPP, R. J. (1964). Parameters affecting the acquisition of Sidman avoidance. *Journal of Experimental Analysis of Behaviour*, **7**, 315–21.

BOLLES, R. C. and GROSSEN, N. E. (1969). Effects of an informational stimulus on the acquisition of avoidance behaviour in rats. *Journal of Comparative and Physiological Psychology*, **68**, 90–9.

HURWITZ, H. M. B. and BOUNDS, W. (1968). Response specification and acquisition of free-operant avoidance behaviour. *Psychological Reports*, **23**, 483–94.

McINTIRE, R. W., DAVIS, H., COHEN, S. I. and FRANCH, E. O. (1968). Sidman avoidance performance under punishment and non-contingent shock conditions. *Psychological Reports*, **22**, 897–903.

SIDMAN, M. (1953). Two temporal parameters of the maintenance of avoidance behaviour by the white rat. *Journal of Comparative and Physiological Psychology*, **46**, 253–61.

SIDMAN, M. (1958). Some notes on "bursts" in free-operant avoidance experiments. *Journal of Experimental Analysis of Behavior*, **1**, 167–72.

SIDMAN, M. (1966). Avoidance Behaviour. In Honig, W. K. (Ed.), *Operant Behaviour: Areas of Research and Application*. New York: Appleton-Century-Crofts.

LIGHT REINFORCEMENT IN THE RAT: THE EFFECTS OF CONTINUOUS AND DISCONTINUOUS PERIODS OF APPARATUS FAMILIARIZATION

D. I. WILLIAMS AND G. LOWE

Department of Psychology, Hull University

The effects of continuous and discontinuous periods of apparatus familiarization on light reinforcement in rats were examined. A previous finding that the light reinforcement effect is greater with longer pretest periods in the dark box was confirmed by comparing 270 min and 30 min periods in the box prior to testing 24 h later. A discontinuous period of 9 daily 30 min sessions, however, produced a bigger effect than the 270 min continuous period. It was concluded that the degree of familiarization with the apparatus is a more important factor than length of time *per se* in the dark in determining the effectiveness of light as a reinforcer.

Introduction

Rats show a significant increase over their normal operant level of bar pressing in a dark box if dim light onset is made contingent upon this response (e.g. Hurwitz, 1956). This effect can be enhanced by a period of pre-test familiarization with the apparatus (Appel and Hurwitz, 1959). Leaton, Symmes and Barry (1963) confirmed this result, but failed to show any greater effect with seven, rather than two, 20 min sessions of prior experience in the experimental box. Lowe and Williams (1968), however, observed a difference between two widely different pre-test periods of 15 min and 10 × 30 min. If a greater light reinforcement effect is found following the longer period of familiarization, is time *per se* the important variable? It is interesting to note that in all the above studies the familiarization periods were of short duration (maximum 30 min), and separated from each other by a 24 h interval. The present study compares the effects of a continuous and discontinuous period in the apparatus prior to test.

Method

Apparatus

The experimental box measured 26 × 26 × 26 cm, and was contained within a sound insulated cabinet. It had a metal mesh floor with walls and lid of Perspex; these were black, except for the end wall which was white. In this end wall was a single Perspex bar, 5 cm wide, 7·5 cm above the floor, extending 2·5 cm into the box. Above the bar an area of the wall 20 × 12 cm could be evenly illuminated from behind. The level of illumination at a point 7·5 cm above the bar and 7·5 cm from the wall was 16 lx. The number of responses in each successive 5 min period was recorded on a print-out counter, and a measure of total response

duration for each rat in each session was also obtained. Six identical sets of apparatus were used simultaneously.

Procedure

Seventy-two male hooded rats (strain PVG/C), aged approximately 100 days old, were used. The test session for each lasted for 60 min in the dark experimental box. For experimental (E) groups, light onset was made contingent upon a response, the light remaining on for as long as the bar was depressed. For control (C) groups the box remained dark when the bar was depressed. Prior to this session each group of 12 rats received different periods of pre-test familiarization in the dark experimental box. Rats in groups E1 and C1 were placed in the dark boxes for 30 min on each of the 9 days prior to the main experimental session. Those in groups E2 and C2 were placed in the dark boxes for 270 min, and in groups E3 and C3 for 30 min on the day prior to testing. Half the animals in groups 2 and 3 were each handled twice on the 8 days prior to their initial experience of the apparatus.

Results

The total numbers of responses made by each group in the main experimental test sessions are shown in Table 1. An analysis of variance on these responses

TABLE I

Total number of responses made by each group during 60 min test period

		Light contingency Response contingent light	No light
Pre-test familiarization period in the dark box	9 × 30 min	609	141
	270 min	393	140
	30 min	250	149

showed a significant difference due to pretest familiarization ($F = 6.65$; $df = 2,66$; $P < 0.01$), a significant difference due to light contingency ($F = 47.78$; $df = 1,66$; $P < 0.001$), and a significant pre-test familiarization × light contingency interaction ($F = 7.21$; $df = 2,66$; $P < 0.01$). From a comparison of individual means, group E1 responded significantly more than group E2 ($F = 9.89$; $df = 1,66$; $P < 0.01$), and group E2 responded significantly more than group E3 ($F = 4.33$; $df = 1,66$; $P < 0.05$). There were no significant differences between those animals in groups 2 and 3 that had been handled prior to testing and those that had not. Mean response durations for E groups were remarkably similar (median 2.39 sec), as were those for C groups (median 1.30 sec). In each case, mean response durations were significantly longer in the E group compared with the corresponding C group (Mann-Whitney U test, $P < 0.01$).

Discussion

Response levels in all no-light control groups were very similar (see Table I), thus differences in the light reinforcement effect following different periods of familiarization in the apparatus cannot be attributed to differences in activity level

resulting from these treatments. Comparison of Group E2 (270 min familiarization) with Group E3 (30 min familiarization) confirms clearly the previous findings (Lowe and Williams, 1968) that responding for light is greater following the longer period of familiarization. The most interesting result, however, is that a greater light reinforcement effect is found after a discontinuous period in the experimental box prior to testing (Group E1), as compared with a continuous period (Group E2). Appropriate control conditions indicated that this difference was due to differential experience of the experimental box, and not due to the additional handling received by Group E1. This result shows clearly that in the case of apparatus familiarization, the "length of time" is not the crucial variable. Hurwitz (1956) suggested that familiarization trials allow competing stimuli to undergo adaptation. This is congruent with the stimulus change hypothesis which assumes that stimulus change is the crucial reinforcing event (cf. Forgays and Levin, 1959), since the more familiar the environment the greater the degree of change resulting from light onset.

Becoming familiar with an environment could be considered as a matter of active investigation on the animal's part. It is clear in this study, from an examination of the number of bar presses made in the no-light pre-test period, that animals in Group E1 spent more time examining the test box than those in Group E2. The mean total number of bar presses in the pre-test period was 117 for Group E1 and 15 for Group E2. An analysis of responding during this time shows a within session decline for all groups; the effect of discontinuous trials is to reinstate investigatory behaviour at the start of each new session. It could be argued then that the environment was more familiar for Group E1 than for Group E2. This would support the conclusion that the more familiar the environment the greater the reinforcing effect of light onset. This is strong evidence for the stimulus change position, and provides apparently insurmountable problems for preference theory. Preference theory (Lockard, 1963, 1966) holds that the reinforcing effect of a stimulus is a function of its preference value as measured in a non-reinforcement situation. Thus the reinforcing effect of light is said to be related to the particular light value used and independent of the environmental context in which light onset occurs. In the present study it can be seen that the same light intensity has a different effect depending upon the animal's pretest experience. This result adds to the increasing body of evidence against preference theory (cf. McCall, 1965, 1966; Williams and Lowe, 1970; Williams, Wells and Lowe, 1971).

References

APPEL, J. B. and HURWITZ, H. M. B. (1959). Studies of light reinforced behaviour: IV. Effects of apparatus familiarization. *Psychological Reports*, 5, 355–6.

FORGAYS, D. G. and LEVIN, H. (1959). Discrimination and reversal learning as a function of change of sensory stimulation. *Journal of Comparative and Physiological Psychology*, 52, 191–4.

HURWITZ, H. M. B. (1956). Conditioned responses in rats reinforced by light. *British Journal of Animal Behaviour*, 4, 31–3.

LEATON, R. N., SYMMES, D. and BARRY, H. (1963). Familiarization with the test apparatus as a factor in the reinforcing effect of change in illumination. *Journal of Psychology*, **55**, 145–51.

LOCKARD, R. B. (1963). Some effects of light on the behaviour of rodents. *Psychological Bulletin*, **60**, 509–29.

LOCKARD, R. B. (1966). Several tests of stimulus-change and preference theory in relation to light-controlled behaviour of rats. *Journal of Comparative and Physiological Psychology*, **62**, 415–26.

LOWE, G. and WILLIAMS, D. I. (1968). Light reinforcement in the rat: The effects of visual pattern and apparatus familiarization. *Animal Behaviour*, **16**, 338–41.

McCALL, R. B. (1965). Stimulus change in light contingent bar pressing. *Journal of Comparative and Physiological Psychology*, **59**, 248–62.

McCALL, R. B. (1966). Initial-consequence-change surface in light contingent bar pressing. *Journal of Comparative and Physiological Psychology*, **62**, 35–42.

WILLIAMS, D. I. and LOWE, G. (1970). Light reinforcement in relation to pre-test illumination. *British Journal of Psychology*, **61**, 379–81.

WILLIAMS, D. I., WELLS, P. A. and LOWE, G. (1971). Light reinforcement, noise, and arousal level. *Nature*. In press.

EFFECTS OF ACETOXYCYCLOHEXIMIDE ON APPETITIVE LEARNING AND MEMORY

DENVER DANIELS†

Department of Psychology, University of Bristol

The effects of direct brain infusions of acetoxycycloheximide (an inhibitor of protein synthesis; ACXH) on acquisition, storage and recall of memory for one-trial appetitive learning were examined in five experiments. ACXH was infused into the rats' hippocampi through implanted cannulas. Control subjects received an equal volume of physiological saline. ACXH was infused (a) 5 h before acquisition, (b) 5 hr before commencement of recall tests, and (c) immediately after acquisition. Each subject's general motor activity was recording during testing. The results indicate that (1) ACXH has similar effects on appetitive and avoidance learning. (2) ACXH administered immediately after acquisition, has no effect on memory. (3) At 4 hr after acquisition memory is affected by ACXH. (4) Short-term memory is unaffected by ACXH and can exist independently of long-term memory. (5) ACXH consistently reduces general motor activity.

Introduction

Several experiments (Barondes and Cohen, 1967, 1968; Cohen and Barondes, 1968; Flexner, Flexner and Roberts, 1967; Daniels, 1971) have provided evidence supporting hypotheses which postulate that cerebral protein synthesis is an essential factor in the storage of long-term memories. In these experiments, acetoxycyclo-heximide (ACXH), a drug which is a powerful inhibitor of protein synthesis (Young, Robinson and Sacktor, 1963), was injected into the brains of mice and rats before and after training in an avoidance learning situation. The results indicated that long-term memory for avoidance learning was severely impaired by ACXH. However, the generality and applicability of these results are limited, since all experiments used punishment in the form of electro-footshock as the reinforcing agent. It is possible that different reinforcers elicit activity in different motivational systems in the brain which are differentially susceptible to ACXH. The learning situation used in the earlier experiments was a Y maze, in which the subjects were required to make either a spatial or a brightness discrimination to avoid footshock. To test the generality of the findings from these experiments, a series of experiments was designed to investigate the effects of ACXH, infused directly into the brain through a double cannula system, on behaviour in a one-trial appetitive learning situation.

The one-trial learning situation used was an adaptation of one used originally by Tenen (1965). The apparatus, procedure and deprivation period devised by

† Now at: Department of Psychology, University of Exeter, Washington Singer Laboratories, Exeter, Devon.

Tenen, were modified (Daniels, 1971b) to allow a more stringent control of motivational factors and permit subjects under the deprivation schedule to be tested daily. The modifications also facilitated measurement of gross body movements as well as of exploratory responses in the apparatus. In the learning situation, following a period of habituation to the apparatus, thirsty rats were required to emit an exploratory response to gain access to water reinforcement. The exploratory response was emitted spontaneously and, by selectively reinforcing this response, the probability of its occurring on a later occasion was increased. One advantage of this learning situation over a Skinner box is that it avoids the long cueing and shaping necessary to produce co-ordinated bar pressing activity, while retaining the advantage of investigating spontaneously emitted behaviour, rather than avoidance behaviour.

Method

Subjects

The subjects were male Lister rats weighing 250–300 g at 90 days of age. All the subjects were housed individually at 80 days of age, with free access to food and water, and maintained on a 12-hr light-dark cycle. The normal growth rate was established as 2 g per day. Following the surgery necessary to implant two cannulae in the brain, the subjects were maintained on ad libitum feeding and drinking until the presurgical growth rate was re-established. On the eighth day after surgery, all subjects were placed on a 23·5-hr water deprivation schedule which was maintained throughout the experiment. Subjects experienced the deprivation schedule for two days before they were introduced to the learning situation.

Surgery

At 90 days of age each subject was deprived of food overnight to facilitate recovery from an injection of Nembutal anaesthetic (42·6 mg/kg) given on the following day. After injection, two stainless steel guide cannulae (Myers *et al.*, 1967) were stereotaxically implanted in the hippocampi of each rat as previously described (Daniels, 1971a). The guide cannulas were covered with protective caps containing stillettes which maintained clear access to the brain tissue. The subjects recuperated in their home cages and were given sufficient time to adapt to the extra weight (1·05 g) on their heads.

Injections

ACXH is an antibiotic drug which inhibits protein synthesis *in vivo* (Young, Robinson and Sacktor, 1963) by preventing the transfer of amino acid from the tRNA (transfer RNA) to the growing polypeptide (Siegel and Sisler, 1963). In this series of experiments ACXH was dissolved in physiological saline in a concentration of 2 mg per 1 ml of saline.

All the experimental subjects in the following series of experiments received 20 μg ACXH in 10 μl physiological saline in each cerebral hemisphere, i.e. a total of 40 μg ACXH in 20 μl saline. All the control subjects similarly received an equal volume of physiological saline. Similar dosages have been used in previous experiments (Daniels, 1970, 1971a).

The injections were made by first removing a protective cap and stillette from the guide cannula and inserting an injector cannula connected by polyethylene tubing to an infusion pump which had been primed with either the drug solution or saline. When inserted fully, the injector cannula protruded 0·5 mm beyond the tip of the guide cannula to ensure maximum infusion of the injected material into the surrounding tissue and minimize the amount of solution which ran up the track of the guide cannula. The infusion pump was programmed to give 10 μl over 10 sec, and then to switch off. This procedure was repeated for the second brain hemisphere.

Histology

When each subject had completed the experiment, a histological assay was made of its brain tissue to verify consistent emplacement of the cannula within the hippocampus. The subjects were sacrificed by overdosing them with Nembutal anaesthetic. The brains were removed and the cerebellum and olfactory bulbs discarded. The remaining brain tissue was frozen quickly by pouring CO_2 syrup (a mixture of dry ice and ethanol) over the brain surface. Later, the brains were sectioned at 40μ intervals on a freezing microtome. Thus correct emplacement was verified in all subjects.

Apparatus

The apparatus shown in Figure 1 consisted of a white aluminium box, 12·5 cm high, 15 cm wide, and 23 cm long. The floor of the box was a steel grid 15 cm × 23 cm. The rectangular box was supported on legs 6·5 cm high, which allowed space for a moving coil receiver beneath the grid floor. The grid floor formed part of a stabilimeter, an addition to Tenen's (1965) original apparatus. In the centre of one of the 23-cm walls, 3·5 cm above the floor, a 3-cm wide, 5-cm high opening was cut to form the opening to an aluminium
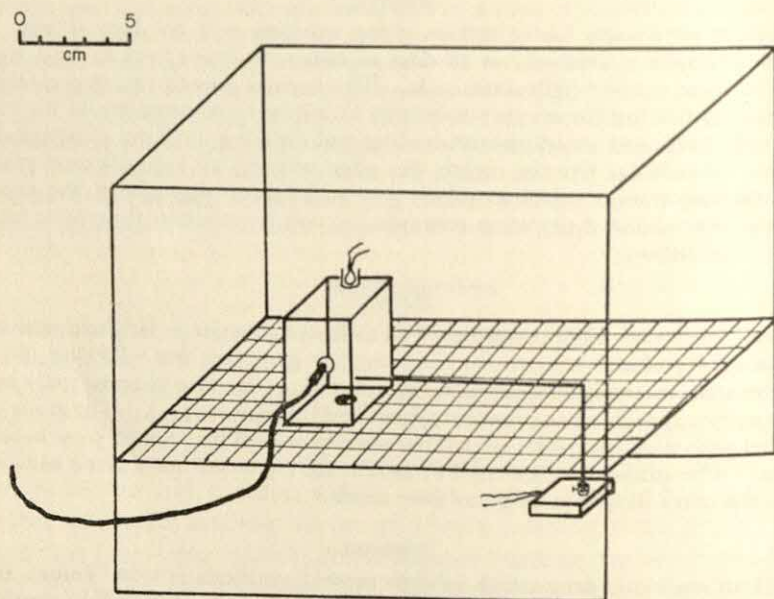
FIGURE 1. The modified Tenen box.

box 3 × 4 × 5 cm. This smaller box was painted black on the inside and was large enough to accommodate a subject's head and the implanted cannulas. A photoelectric cell and opposing light bulb (2·2 V) were placed on the smaller box, 1 cm from the entrance. Interruption of this light beam activated a relay counter which recorded the event. In this way a cumulative record was acquired of the number of times a subject's head entered the smaller box during the time the subject was in the apparatus. Centred in the rear wall of the smaller box, was a chrome-plated water valve, which could be connected by polyethylene tubing to a water supply. The grid floor was made from steel wire with small springs welded to the four corners on the underside. The springs rested lightly on two 1-cm wide shelves, riveted to the inside of each 23-cm wall. A length of stainless steel wire (13·5 cm) was welded to the centre of the grid floor and, at 8·5 cm along its length, bent at a right angle. The protruding 5 cm was fed into a moving-coil receiver, thus any movement of the floor was transmitted to the moving coil, as in a gramophone pick-up.

The moving coil produced a discrete signal in response to movement of the grid floor and this signal was fed into an amplifier to operate a relay counter. A cumulative record of gross body movements was thus obtained during the time the subject was in the apparatus.

The apparatus was covered with a removable white perspex lid. During training and testing, the apparatus was placed in a dark, sound-protected room. The recording equipment was kept outside the sound-protected room.

General procedure

When subjects had completed 2 days on the 23·5-hr water deprivation schedule, they were introduced to the apparatus. This was done 30 min before the normal 30-min period of access to water in the home cage. Although the water valve was a fixture in the apparatus, water was only available from it on the fifth day of the experiment.

On Day 1 of the experiment a subject was placed in the apparatus in the sound-protected and darkened room, and the lid was secured. The subject remained in the apparatus for 4 min; this period was timed on a stop watch. During the 4-min period the subject's general activity (GA) was recorded cumulatively by means of the stabilimeter. The number of times the subject thrust its head into the smaller box (hole exploratory activity: HE) was also cumulatively recorded. After 4 min the data from the recorders was noted and the subject returned to its home cage. The recorders were zeroed in preparation for the following subject. This procedure was repeated, at the same time of day, on Days 2, 3 and 4 of the experiment. In this way, a baseline level was achieved for the two behavioural indices, HE and GA.

Both HE and GA decreased markedly over the first three days of the experiment. However, as will be seen from the results of Experiment I, the Day 4 scores on GA and HE reflect a relatively stable level of responding. This period, Days 1–4, of the experiment was considered a period of habituation to the apparatus.

On Day 5 of the experiment the subjects were placed in the apparatus at the usual time, but water was available at the water valve in the smaller black box. Once the subjects had located the water they were allowed to drink for 10 sec before being removed to their home cages. This phase of the experiment was called acquisition. HE and GA were not recorded on Day 5.

Recall tests were commenced 24 hr after acquisition, on Day 6 when once more each subject was placed in the apparatus for 4 min. Water was not available at the water valve during the recall tests, and HE and GA were again recorded. Recall tests were made at 24 hr intervals from Day 6 to Day 9. Since no reinforcement was given during the recall tests, this period could be regarded as a period of extinction.

Computation

Throughout the experiments, and for each subject, the HE and GA scores obtained on Day 4 were subtracted from HE and GA scores recorded during the subsequent tests, and a statistical analysis made of the resulting scores. The tests used, except where indicated in the text, were the t-test and the F test for homogeneity of variance (Winer, 1962).

Experiment I

The effect of ACXH on behaviour in a one-trial appetitive learning situation

Procedure

Twelve subjects were surgically prepared, deprived of water, and habituated as described.

Results

The results of the statistical analysis are given in Table I.

In Figure 2 the mean HE and GA scores are given for each group throughout the experiment. It can be seen from Figure 2 that there was a marked decrease

TABLE I

Statistical analysis of Day 4 weighted data from Days 6 to 9 of Experiment I

Day	Mean$_e$	S.D.$_e$	Mean$_c$	S.D.$_c$	t	F
			Hole explorations			
6	−3·1	4·9	8·1	5·8	3·72†	1·3
7	−3·8	6·4	3·16	5·9	2·06	1·5
8	−3·16	6·1	0·83	3·5	1·38	2·97
9	−4·0	6·66	−1·0	3·34	0·98	3·96

Day	Mean$_e$	S.D $_e$	Mean$_c$	S.D.$_c$	t	F
			General activity			
6	−0·13	0·3	44·5	19·3	5·65	3153·52†
7	−0·28	0·46	1·7	1·05	4·21†	5·09
8	—	—	—	—	—	—
9	—	—	—	—	—	—

† Significant at $P = 0.01$; e = experimental group; c = control group.
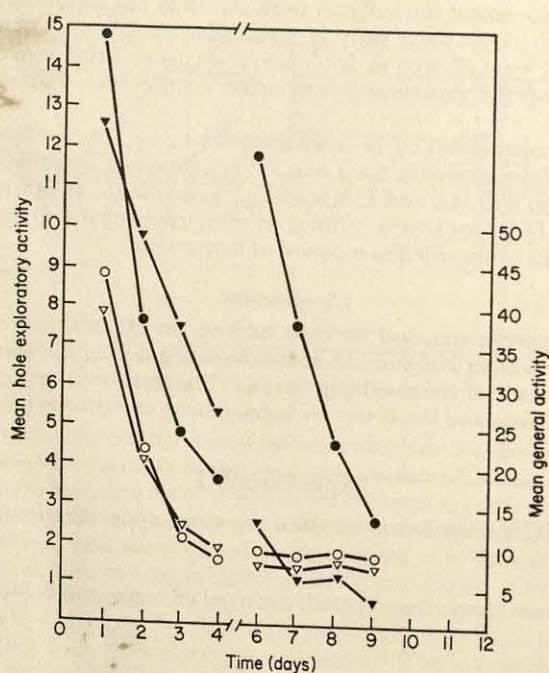
FIGURE 2. ACXH/Saline infused into the brain 5 h before acquisition; subjects tested for recall at 24-h intervals. —●— HE saline group. —▼— HE ACXH group. —O— GA saline group. —▽— GA ACXH group.

in HE by the experimental group, but the statistical analysis of the Day 4 weighted data from Days 6 to 9 shows that the differences are only significant on Day 6. The GA scores were only significantly different on Day 7.

GA appeared to be relatively unaffected by ACXH compared with HE. The experimental subjects explored the small box from which they had received water on Day 5, significantly less often than the control subjects, and behaved as if they had not received the reinforcement. It is unclear from these results whether ACXH achieved its effect on HE by preventing acquisition or storage, or by blocking recall of memory for this particular task.

Experiment II

The effect of ACXH on behaviour during the recall of memory

Procedure

Thirteen subjects were surgically prepared as described, and the water deprivation schedule was imposed. The procedure for the first 4 days of the experiment was similar to that described in Experiment I. Seven subjects were included in the saline-injected control group, and six in the drug-injected, experimental group. The intracerebral infusions were made on Day 6, 5 hr before the recall test made at 24 hr after acquisition.

Results

A summary of the statistical analysis of each subject's weighted scores and of the main changes in HE and GA, is given in Table II. The data in Table II show that there were no significant differences in HE for experimental and control groups over the first two days after acquisition or on the fourth day after acquisition, but Day 8 HE scores (i.e. HE scores on the third day after acquisition) were

TABLE II

Summary of a statistical analysis of the Day 4 weighted data over Days 6 to 9 from Experiment II

Day	Mean$_e$	S.D.$_e$	Mean$_c$	S.D.$_c$	t	F
			Hole explorations			
6	8·66	3·7	10·71	3·94	0·95	1·12
7	6·0	3·6	3·57	2·37	1·44	2·3
8	3·0	1·67	−0·42	1·9	3·42†	1·29
9	−0·33	1·36	−0·14	1·21	0·26	1·26
			General activity			
6	−1·66	0·86	0·28	0·95	3·9†	1·35
7	0	0·89	0	0·57	0	2·4
8	−0·5	1·04	0·14	1·3	0·94	1·6
9	−0·5	0·54	0·28	0·48	2·7	1·26

† Significant at $P = 0.001$; e = experimental group; c = control group.

significantly different. GA was only significantly affected on Day 6. The latter finding indicated that ACXH administered 5 hr before the recall test reduced general motor activity in rats, and this could have been a contributory factor to the impairment observed in HE activity in Experiment I.

Experiment III

The effect of ACXH on memory when administered immediately after acquisition

The one-trial learning situation, unlike a maze-learning situation, allows a distinction to be made between the acquisition and a conceived period of storage of memory. It provided an opportunity to evaluate the effect of ACXH on the storage of memory following a relatively discrete acquisition trial.

Procedure

Fourteen subjects were prepared surgically as described, water deprived and habituated to the apparatus over the first four days of the experiment. Within 30 sec of being removed from the apparatus, following the acquisition trial on Day 5, the experimental group received bitemporal infusions of ACXH and the control group received physiological saline. After treatment, each subject was returned to its home cage. Recall tests were made at 24 hr intervals commencing on Day 6.

Results

The results are given in Table III. Table III shows the result of the statistical analysis on the Day 4 weighted scores over Days 6, 7, 8 and 9 for HE and GA. No significant differences were detected in either HE or GA scores. Despite the relatively discrete acquisition trial in this learning situation, ACXH administered

TABLE III

Statistical evaluation of HE and GA scores over Days 6 to 9 in Experiment III after they had been weighted by Day 4 scores

Day	Mean$_e$	s.d.$_e$	Mean$_c$	s.d.$_c$	t	F
			Hole explorations			
6	6·0	2·4	8·57	2·07	2·12	1·4
7	3·7	1·49	4·14	1·34	0·56	1·23
8	0·14	1·06	2·14	2·26	2·11	4·5
9	−0·85	1·06	0	1·0	1·54	1·14

Day	Mean$_e$	s.d.$_e$	Mean$_c$	s.d.$_c$	t	F
			General activity			
6	0·14	0·89	0·85	0·69	1·66	1·7
7	−0·28	0·48	0·28	0·75	1·6	2·4
8	0·14	0·69	0·14	0·89	0	1·7
9	−0·7	0·48	0·42	0·7	0·81	2·6

e = experimental group; c = control group.

immediately after acquisition had no effect on memory at 24 hr or over the subsequent 3 days. Thus the conceived period of storage following acquisition is not affected by ACXH.

However, although it has been shown that ACXH does not prevent acquisition of shock-motivated Y-maze performance (Daniels, 1971a), a similar finding remained to be demonstrated in an appetitive situation.

Experiment IV

Effects of ACXH on short and long term memory

The results from previous experiments using ACXH in mice (Barondes and Cohen, 1967), and rats (Daniels, 1971a) have indicated that there are at least two memory systems, one susceptible and one insusceptible to ACXH. When ACXH was intracerebrally infused into rats 5 hr before acquisition (Daniels, 1971a), memory for avoidance learning at 3 hr after acquisition was unimpaired whereas memory at 6 hr, 24 hr and 7 days after acquisition was severely impaired. It was reasoned therefore, that if subjects under the influence of a pre-acquisition
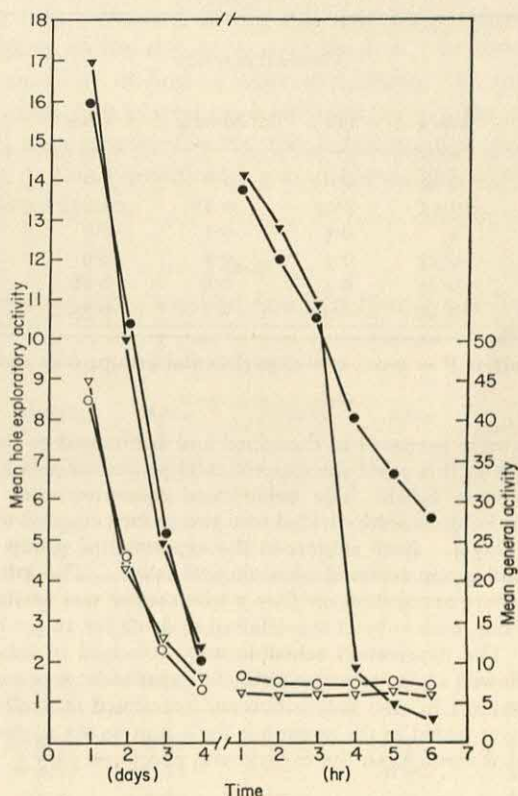
FIGURE 3. ACXH/Saline given 5 h before acquisition. Recall tests made at hourly intervals after acquisition. —●— HE saline group. —▼— HE ACXH group. —○— GA saline group. —▽— GA ACXH group.

infusion of ACXH, in the present learning situation, could express memory over the first 5 hr following acquisition, then it could be concluded that ACXH did not block acquisition. Further, by testing the subjects for recall at hourly intervals after acquisition, it might be possible to determine the time at which memory becomes dependent on cerebral protein synthesis.

TABLE IV

Statistical evaluation of Day 4 weighted data from hours 1–6 in Experiment IV

Hour after acquisition	Hole explorations					
	$Mean_e$	S.D._e	$Mean_c$	S.D._c	t	F
1	11·7	2·1	11·8	1·2	0·13	2·7
2	10·3	1·6	10·4	1·13	0·16	2·15
3	8·6	0·86	8·6	1·41	0	2·6
4	−0·4	1·23	6·1	1·6	9·6†	1·7
5	−0·7	1·09	4·6	1·65	8·2†	2·3
6	−0·66	0·86	3·5	1·74	6·51†	4·03

Hour after acquisition	General activity					
	$Mean_e$	S.D._e	$Mean_c$	S.D._c	t	F
1	−0·88	0·6	0·6	1·0	4·0†	2·7
2	−0·55	1·01	0·88	0·6	3·67†	2·8
3	0	0·5	0·1	0·6	0·42	1·4
4	−0·33	0·5	0·1	0·9	1·26	3·4
5	−0·55	0·5	0·6	0·86	3·61†	2·7
6	−0·55	0·52	0·1	0·33	2·13	2·5

† Significant at $P = 0.01$; e = experimental group; c = control group.

Procedure

Eighteen subjects were prepared as described and habituated to the apparatus over the four day period. Up to this point the experimental procedure was no different from that followed in Experiments I–III. The behavioural measures were again HE and GA. Following Day 4, the subjects were divided into two groups equated on the basis of the HE scores recorded on Day 4. Each subject in the experimental group received ACXH and subjects in the control group received physiological saline. The infusions were made via the cannulae 5 hr before acquisition on Day 5 when water was available in the apparatus. After locating the water, each subject was allowed to drink for 10 sec before being returned to the home cage. The deprivation schedule was prolonged in this experiment and the subjects were not allowed access to water until the experiment was completed. The recall tests were commenced at 1 hr after acquisition and continued thereafter at hourly intervals. Thus, each subject was tested in the apparatus for 4 min on six successive occasions. HE and GA were recorded throughout the experiment, except on Day 5.

Results

The results are given in Figure 3 and Table IV. Table IV shows a summary of the statistical analysis of HE and GA scores over hours 1–6. Figure 6 shows

the mean raw score for HE and GA in both groups throughout the experiment. Although over the first 3 hr following acquisition the HE scores of the experimental and control groups were not significantly different, the GA scores in the experimental group were consistently and significantly lower than those of the control group (see Table IV). Later GA score differences were not significant. At hours 4, 5 and 6 after acquisition, the HE scores of the experimental subjects were markedly (see Fig. 3) and significantly (see Table IV) different from the control group.

Experiment V

A further evaluation of ACXH on learning and memory

It was possible that the results of Experiment IV were due to a more rapid re-habituation, or to a faster rate of extinction in the experimental group. Either of these possibilities could account for the sharp drop in responding at hour 4 in Experiment IV, since all the subjects in that experiment were exposed to the apparatus at regular and repeated intervals without reinforcement. However, if faster re-habituation and/or more rapid extinction were really the cause(s) for the apparent loss of memory at hour 4 in the experimental group, then limiting the exposure of the subjects to the apparatus over the first 3 hr should result in both groups responding similarly at hour 4 after acquisition. If, on the other hand, re-habituation and extinction played no significant part in the observed decrease in HE, and ACXH was responsible for the behavioural impairment, then the level of responding in the experimental group should be significantly different from that of the control group.

TABLE V

Statistical analysis of Day 4 weighted data from hours 4–7 in Experiment V

Hour after acquisition	Hole explorations					
	Mean$_e$	S.D.$_e$	Mean$_c$	S.D.$_c$	t	F
4	−0·16	0·75	10·66	3	8·6	16·0†
5	−0·16	0·4	7·6	2·06	9·11	25·6†
6	−0·33	0·51	7·33	2·06	8·82	16·0†
7	—	—	—	—	—	—

Hour after acquisition	General activity					
	Mean$_e$	S.D.$_e$	Mean$_c$	S.D.$_c$	t	F
4	−0·16	0·98	0·5	0·54	1·4	3·2
5	−0·66	0·51	0·5	0·54	3·7†	1·125
6	−0·5	0·54	0·33	1·03	1·74	3·55
7	−0·83	1·72	0·33	0·51	0·68	11·125

† Significant at $P = 0·01$; e = experimental group; c = control group.

Procedure

The procedure in this experiment, up to and including the acquisition trial on Day 5 was similar to that described for Experiment 4. Each group contained 6 subjects. The only change in procedure was introduced after acquisition when, following the prolongation of the deprivation period, the first recall test was made at hour 4 and subsequent recall tests at hourly intervals, up to hour 7. For the 4 hr following acquisition, the subjects remained in their home cages.

Results

A summary of the statistical analysis of Day 4 weighted data from hours 4 to 7, is given in Table V. Due to the values calculated for F, in a test of the homogeneity of the variance, being significant, it was not possible to attribute the large value calculated for t with statistical significance. However, the mean raw scores plotted as a graph in Figure 4 show marked differences in the HE scores of both groups.

General Discussion

The results of Experiment I showed that when ACXH was administered 5 hr before acquisition, memory for appetitive learning at 24 hr is severely impaired. This result is similar to those of previous experiments using ACXH in an avoidance learning paradigm (Barondes and Cohen, 1967; Daniels, 1971a). Further, the present results show that the effect is maintained over at least 4 days. If different reinforcement/motivational systems exist for avoidance and appetitive learning then they appear to react in similar ways to ACXH.

However, from the results of Experiment I the possibility remained that ACXH could have achieved its effect on memory by blocking acquisition, or recall, or by preventing the storage of memory. These possibilities were examined in Experiments II–V.

The results of Experiment II support and extend the findings from maze experiments with rats (Daniels, 1971a) which indicated that ACXH had no effect on the recall of memory for an avoidance task. Here it was shown that ACXH does not prevent recall of memory for an appetitive task at 24, 48 and 96 hr after acquisition, but does affect GA 5 hr after injection. Experiment II therefore indicated that ACXH did not achieve its effect on memory by blocking a recall mechanism operating at the time of recall.

When ACXH was administered immediately following the acquisition trial (Experiment III) no effect was observed on subsequent behaviour. Thus, it appeared that in this appetitive learning situation, as in the avoidance-learning studies (Barondes and Cohen, 1967; Cohen and Barondes, 1968; Daniels, 1971a), ACXH must be injected into the brain some time *before* acquisition for long-term memory to be affected. In other words ACXH must be present in the brain at the time of acquisition for long-term memory to be impaired.

The data from Experiment IV indicate that ACXH caused a decrease in GA over the first 2 hr following acquisition (see Fig. 3 and Table IV) but this effect did not interfere with the subjects' ability to recall over the first 3 hr. It is also evident from this experiment, that ACXH did not achieve its effect on memory by preventing acquisition, since memory over the first 3 hr (short-term memory) was

unimpaired. This finding is consistent with previous findings in avoidance learning (Barondes and Cohen, 1967; Cohen and Barondes, 1968; Daniels, 1971a), and suggests that short-term memory may not be dependent on continuing protein synthesis in the brain, and can exist in the absence of long-term memory. Reference to Figure 3 shows that HE scores fell significantly at hours 4, 5 and 6, and this is interpreted as indicating that between 3 and 4 hr after acquisition, memory becomes dependent on protein synthesis in the brain, and supports previous findings (Daniels, 1971b).
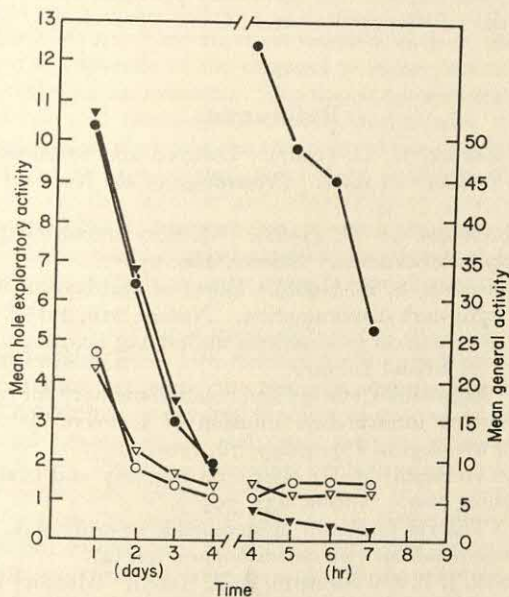
FIGURE 4. ACXH/Saline given 5 h before acquisition and hourly recall tests delayed until 4 h after acquisition. —●— HE saline group. —▼— HE ACXH group. —○— GA saline group. —▽— GA ACXH group.

The results of Experiment V indicated that the experimental group exhibited a gross memory impairment on the first recall test at hour 4, compared with the control group. It seems likely therefore, that the behaviour observed over hours 1 to 6 in Experiment IV were not due to rehabituation, or to more rapid extinction in the experimental group. This result is interpreted as indicating that ACXH interfered with a system or systems in the brain which mediate long-term memory. Since ACXH is known to interfere with protein synthesis, the latter is likely to be involved in the establishment and maintenance of long-term memory.

In sum, it appears that ACXH must be administered intracerebrally 5 hr prior to acquisition, for long-term memory in this situation to be impaired. This effect may be the result of an interference with:

(a) a system necessary for the *storage* of long-term memory, activity in which is initiated at the time of acquisition (since ACXH given after acquisition does not affect long-term memory);

(b) a system for the *recall* of long-term memory which is established during acquisition (ACXH given prior to recall has no effect on memory).

On the other hand, ACXH does not appear to effect:

(i) acquisition of long-term memory;
(ii) recall of long-term memory when given 5 hr before the recall test;
(iii) storage of memory when given immediately after acquisition.

References

BARONDES, S. H. and COHEN, H. D. (1967). Delayed and sustained effects of acetoxycycloheximide on memory in mice. *Proceedings of the National Academy of Science,* **58,** 157–64.

BARONDES, S. H. and COHEN, H. D. (1968). Memory impairment after subcutaneous injection of acetoxycycloheximide. *Science,* **160,** 556–7.

COHEN, H. D. and BARONDES, S. H. (1968). Effect of acetoxycycloheximide on learning and memory of a light-dark discrimination. *Nature,* **218,** 271–3.

DANIELS, D. (1970). Biochemical mechanisms underlying learning and memory. Ph.D. Thesis, University of Bristol Library.

DANIELS, D. (1971a). Acquisition, storage and recall of memory for brightness discrimination by rats, following intracerebral infusion of acetoxycycloheximide. *Journal of Comparative and Physiological Psychology,* **76,** 110–8.

DANIELS, D. (1971b). Actinomycin-D: effects on memory and brain RNA synthesis in an appetitive learning task. *Nature,* **231,** 395–7.

DEGROOT, J. (1959). The rat forebrain in stereotaxic co-ordinates. *Verhandelingen van het K. nederlandsche akademie van wetenschappen,* **52,** (4).

FLEXNER, L. B., FLEXNER, J. B. and ROBERTS, R. B. (1967). Memory in mice analysed with antibiotics. *Science,* **155,** 1377–83.

HYDÉN, H. (1966). Introductory remarks to the session on memory processes. In SCHMITT, F. O. and MELNECHICK, T. (Eds.), *Neurosciences Research Symposium Summaries,* Vol. I. Pp. 278–93. M.I.T. Press.

JOHN, E. R. (1967). *Mechanisms of Memory.* New York: Academic Press.

MYERS, R. D., CASADAY, G. and HOLMAN, R. B. (1967). A simplified intracranial cannula for chemical stimulation or long-term infusion of the brain. *Physiology and Behaviour,* **2,** 87–8.

ROBERTS, R. B. and FLEXNER, L. B. (1969). The biochemical basis of long-term memory. *Quarterly Review of Biochemistry and Biophysics,* **2,** 135–73.

SIEGEL, M. R. and SISLER, H. D. (1963). Inhibition of protein synthesis in vitro by cycloheximide. *Nature,* **200,** 675–6.

TENEN, S. S. (1965). Retrograde amnesia from electroconvulsive shock in a one trial appetitive learning task. *Science,* **148,** 1248.

WINER, B. J. (1962). *Statistical principles in experimental design.* New York: McGraw Hill.

YOUNG, C. W., ROBINSON, P. F. and SACKTOR, B. (1963). Inhibition of the synthesis of protein in intact animals by acetoxycycloheximide and a metabolic derangement concomitant with this blockade. *Biochemical Pharmacology,* **12,** 855–65.

BOOK REVIEWS

MILNER, P. M. *Physiological Psychology.* New York: Holt, Rinehart and Winston. 1970. Pp. 531. £6·10.

This is a textbook for students that already have some background in experimental psychology but do not necessarily have any prior knowledge of physiology or anatomy. The book is well organized; there are six main sections, each of which is prefaced by an introduction showing the relevance of the chapters to come, and each of the 21 chapters ends with a brief summary of its contents. The first three sections, entitled Foundations (which includes anatomy and techniques), Motor Mechanisms, and Sensory Systems, occupy something over 50% of the book; the allocation of less than half the space to central mechanisms is traditional in physiological psychology textbooks, but is not necessarily desirable. The chapter on the anatomy and physiology of neurons occupies 13 pages, whereas that on the frontal lobes occupies 9 pages; if young psychologists are to be persuaded that physiological techniques may usefully be used in the assault on topics of psychological interest, surely relatively more space must be given to those areas which seem most directly relevant to such topics.

The first three sections are the most successful part of the book; they are simply and clearly written, and provide a good introduction for the students at whom the book is aimed (the physiology of the neuron is, for example, described without reference to the Nernst equation). The section on sensory systems is not, as in some textbooks, over-indulgent to vision and hearing—taste and smell and the somato-sensory system also obtain reasonable coverage. There is a consistent effort to relate sensory data to perceptual phenomena, and the perceptual effects of brain lesions in humans are reviewed. The treatment of sensory systems is, however, only at an introductory level, and does not in itself provide sufficient information on certain important issues. For example, dark adaptation receives only superficial coverage, in which Rushton's work on retinal adaptation pools is not mentioned; similarly, sound localization is discussed without reference to the importance of wavelength, or to the relevant psycho-physical data.

The final three sections are entitled Alerting Mechanisms (containing clear reviews of the physiological correlates of sleep and attention), Motivation and Emotion, and Memory and Learning. The section on Motivation and Learning has chapters on the basic drives, and on such topics as activity, curiosity and fear-motivated behaviour. It also contains chapters on self-stimulation and the frontal lobes, which are the least successful in the book, because, in each case, of a failure to deal adequately with theoretical issues. There is no attempt, in the chapter on self-stimulation, to discuss the relevance of the phenomenon to traditional psychological theories of reinforcement (e.g. to drive-reduction theory); in place of this, Milner devotes a disproportionate amount of space to elaborating, not very clearly, his own theory of learning, which seems to be derived basically from Tolman, with additions inspired by Hebb and Deutsch. The resulting theory is vague and would be distinctly unhelpful to introductory students. The chapter on the frontal lobes does not succeed either in systematizing the observed deficits, or in showing what sort of explanations should be sought, and what implications such explanations, if successful, might have.

There are two chapters in the section on Memory and Learning, the second of which is a properly sceptical account of the efforts to date to elaborate biochemical theories of memory storage. The first chapter, although entitled Physiological Contributions to Learning Theory, is in fact about theories of neural consolidation, and the distinction between long and short term memory. If physiological psychology has contributed to learning theory (in the conventional sense of the term), then Milner does not show us

how or where; the weakness of the sections on learning is precisely his inability to bridge the gap between physiological findings and behavioural theories.

The first half of this book, then, provides an introduction to the peripheral aspects of the nervous system that is clear, if sometimes too elementary, and is particularly suitable for students without a biological background; the second half is less readable, primarily due to the failure to show the general relevance of many of the topics discussed. It is fair to add, however, that this weakness is shared by its major textbook rivals.

EUAN M. MACPHAIL

BAKKER, D. J. and SATZ, P. (Eds). *Specific Reading Disability: Advances in Theory and Method.* Rotterdam: University Press, 1970. Pp. 166. D.Fl. 35.00.

This is a timely book. It shows that research on reading disability really can be worthwhile when it draws its inspiration directly from neurology and experimental psychology. It may thus lead to some of the burgeoning interest in cerebral dominance being channelled into the study of reading backwardness. Although we do not yet have a detailed understanding of the processes of maturation and learning that underlie development of language and manual skill, there are grounds for optimism in recent progress. No longer should anyone expect clear results to emerge from his research unless he separates consistent sinistrals from individuals with mixed handedness; no longer should we expect the same one of a number of more basic performances to correlate highest with reading ability *at all ages.*

The book sprang from a symposium held in 1968, but includes a fair number of more recent references. The eight contributions, particularly two very useful reviews by Satz and Sparrow and by Marian Annett on dyslexia and handedness respectively, are written for a non-specialist audience.

MARK HAGGARD

FJERDINGSTAD, EJNAR J. (Ed.). *Chemical Transfer of Learned Information.* Amsterdam and London: North-Holland, 1971. Pp. xxvii + 268. £7.90.

It *can't* exist, say some. Others, whose scepticism is rational, but who have yet to evaluate the evidence for themselves, may care to know that the first question posed by the editor of this book is whether there really is a phenomenon. He is to be congratulated on the collection of sober answers gathered from 16 major contributors to the field—mostly believers but not all, as is appropriate. Other reviews to date have a single author or are much longer multi-author jobs, often including lengthy flights of fancy.

Various hormones affect behaviour, so one should not be very surprised that brain extracts can do so too. Dyal finds, up to 1970, 133 reported experiments with transfer effects having two-tailed p less than 0.05, and 130 null or equivocal results. He claims, as he has every right to do, that positive results mean more than negative ones, given our ignorance of possible mechanisms by which injected preparations from the brains of trained animals could affect behaviour.

The next question is whether *information* can be transferred chemically. A "cross-over" experiment is the minimally adequate design to establish this: extracts from one situation facilitate behaviour in that situation but not in a second situation, and (in the same experiment) conversely. Such a design does not of course distinguish between transfer of associative factors and transfer of stimulus or response bias. That simply raises the next question—what information can be transferred? Some authors in this book object to the crossover design because of this ambiguity and advocate the use of transfer of discrimination training as an alternative. It is not an alternative, but introduces an additional factor which must be assessed within a crossover design. A positive result in such a design has apparently yet to be obtained. That is, there is no evidence yet that information of a describable sort has been transferred. Some of the authors seem unnecessarily concerned to demonstrate the phenomenon for "learning" in a sense which excludes habituation and other

unreinforced or non-associative instruction by experience. That would be fine, unless the phenomenon exists only for stimulus reactivity bias and the like. The only crossover experiments mentioned in the book, two by Ungar, do not have test procedures that can specify the stimulus information, the response information or the associative information that might be transferred. Nevertheless, one of these results is an intramodal transfer of habituation without intermodal transfer.

It is obviously very risky to try and isolate a transfer factor from the brain extract without using the crossover design as the "assay" for activity. Ungar has isolated and structurally defined a peptide which keeps mice in a bright compartment when a dark compartment is available ("scotophobin"). This remarkable achievement will not establish the existence of information transfer until we also have a "scotophilin" or a "photophobin", and a successful test of both peptides in a crossover design.

Then, and only then, do other fascinating questions arise. What do the transfer factors do to the brain? Have the transfer effects anything to do with personality development or with memory retention? Or are they a nonfunctional artefact deriving from the degree of chemical specificity in neural systems that some workers on regeneration have postulated?

<div align="right">D. A. BOOTH</div>

GAZZANIGA, M. S. *The Bisected Brain*. New York: Appleton-Century-Crofts. 1970. Pp. 172. $12.10.

"A curious aspect of split-brain research is the oblique way it speaks to any particular theory of brain function. The startling phenomenon itself offers little insight into the mysteries of mind and brain in terms of illuminating real neural mechanisms. . . . Certainly there are few, if any, propositions in neurobiology that would be rejected or accepted on the basis of the original split-brain experiments" (p. 141).

One cannot help but agree with this cautious statement by Gazzaniga, one of the leading practitioners of split-brain work. Most such research has served merely to confirm what we knew already from other evidence about laterality in the brain, and the excitement that surrounded the initial use of the split-brain preparation has not been justified by the results obtained with it since. Even the finding that the two disconnected hemispheres act independently is of less interest to scientists than to theologians whose professional concern is with the soul rather than the brain; had the opposite result been obtained it would have been most interesting and puzzling.

The present volume is an excellent review of the research to date. It gives a particularly good account of the ingenious detective work that has gone into uncovering the devious ways in which the two split hemispheres learn to signal to one another (often through peripheral movements of muscles sensed by both hemispheres) when the experimenter sets a task that demands their co-operation. Considerable space is devoted to describing techniques, both behavioural and surgical, and little previous knowledge is assumed so that the book can be read with profit by undergraduate students.

Although the writing is always clear, the author has been badly served by his publishers. There is an excessively large number of misprints, the English is clumsy in places and the book's value as a source of reference is greatly reduced by the omission of a name index.

<div align="right">N. S. SUTHERLAND</div>

MROSOVSKY, N. *Hibernation and the Hypothalamus*. New York: Appleton-Century-Crofts, 1971. Pp. 287. $16.00.

This book is on a specialized, and perhaps even rather a recondite subject. Unlike many authors of monographs, however, Dr Mrosovsky does not ramble drearily along mentioning all the papers he has ever read in his chosen field; he pursues an argument. The argument is by no means watertight as he is the first to admit, but how much less compelling is the cut and dried than the realm of well marshalled speculation.

The author proposes that hibernation is a cunning and sophisticated biological solution

to the problem of severe annual climatic changes. Unlike the other great solution—migration—hibernation depends heavily on physiological adaptations as well as behavioural ones. Nor are the physiological mechanisms primitive: hibernators become cool and torpid during the winter not because like cold blooded animals they have no physiological means for any alternative, but because they must conserve energy during the long cold winter when food is absent and the environment hostile. The pattern of hibernation thus involves eating voraciously during the autumn to stock up with bodily energy reserves and then turning down metabolism to a very low tick-over for several months.

Dr Mrosovsky is concerned to show that the annual periodicity of hyperphagia and aphagia, and the variations of body temperature are made possible by ordinary hypothalamic mechanisms present in all mammals and that hibernation is a case of convergent evolution. He argues that the lesions of the hypothalamus which result in disturbances of feeding behaviour, disrupt function in ways that accidently reveal the capacities of the hypothalamus of non-hibernators to induce voracity or abstention from food.

One interesting aspects of this work is its relationship with the lesion/stimulation school of physiological psychology. The brain provides the same invitation to open it up and take it apart as does a clockwork toy for a school-boy. This examination has not enabled the psychologist to discover how structure gives rise to function, partly because the brain is not made of clockwork. Instead, only when a pathological syndrome produced by a direct manipulation of the brain (such as the finickiness and disinclination to work of hypothalamic hyperphagic animals) can be observed in some aspects of normal behaviour (i.e. in animals preparing for hibernation) does the pathological behaviour acquire any real interest. This phenomenon of lesion- or stimulation-induced pathology of the hypothalamus making sense insofar as it is mirrored somewhere by normal behaviour occurs not only in Dr Mrosovsky's work. It has recently been pointed out that there is a parallel between recovery from aphagia and the normal development of eating patterns in the new-born, that the oddities attending self-stimulation behaviour are essentially those of high incentives with low internal motivation and zero delay of reinforcement, and that the over-reactivity to external stimuli of fat rats with holes in their brains is similar to the behaviour of fat people with intact brains. What the lesion or stimulation has achieved in each case has been not to reveal the inner workings of the nervous system, but to draw attention to a particular aspect of normal behaviour. Because the normal behaviour is normal it is (one feels) a better proposition for satisfactory explanation.

Our understanding of the normal mechanisms underlying hibernation is so far necessarily limited. One problem for instance, is to determine the nature of the annual cycles of behaviour and metabolism. Dr Mrosovsky devotes a chapter to discussing the notion that hibernation (at least in some species) is controlled by circannian oscillators analogous to the much better understood circadian oscillators, tentatively concluding (I think wrongly) that it probably is not. The difficulty as he points out is that relevant information is just not available, often because of the time it takes to collect data on annual cycles.

Other hypotheses to which Dr Mrosovsky gives an airing are that the changes of body temperature accompanying hibernation are under the control of a thermostat with a sliding set point, and Strumwasser's idea that the series of coolings and rewarmings of the body preceding complete hibernation are "test-drops" which serve to check out and accustom the system for its forthcoming winter rigours.

The book also contains a wide selection of facts and many of these will be of interest to those who do not know much about hibernation. Nevertheless this book is not for those who revere the fact as the pinnacle of scientific achievement. Quite possibly "some minds may be exasperated by the lack of attempts to check directly whether hypothalamic systems concerned with obesity and temperature are actually changing in hibernators" (p. 216). Evidently my mind is not of that stamp. I found that the book stimulated my interest in the subject. It is also written with some well-turned phrases (how much better occasionally to read of an animal becoming "distinctly corpulent" rather than always "obese"), and has been put together with considerable care.

KEITH OATLEY

Publications Received

AIKEN, L. R. Jr. *Psychological and Educational Testing*. Boston: Allyn & Bacon. 1971. Pp. 346. $9.50.

ATKINSON, R. C. (Ed.). *Contemporary Psychology: Readings from Scientific American*. San Francisco: W. H. Freeman. 1971. Pp. 484. £2·50.

BAGLEY, C. *The Social Psychology of the Epileptic Child*. University of Miami Press. 1971. Pp. 307. $12.00.

BANNISTER, D. AND FRANSELLA, F. *Inquiring Man: the Theory of Personal Constructs*. Harmondsworth: Penguin Books Ltd. 1971. Pp. 221. £0·50.

BARBER, T., DiCARA, L. V., KAMIYA, J., MILLER, N. E. AND STOYVA, J. (Eds). *Biofeedback and Self-control*. 1970. Chicago: Aldine. 1971. $12·00.

BINDRA, D. AND STEWART, J. *Motivation*. Harmondsworth: Penguin Books Ltd. 1971. Pp. 529. £0·75.

BLOOM, L. *Language Development: Form and Function in Emerging Grammars*. M.I.T. Press. 1970. Pp. 270. £4·20.

BLUMENTHAL, A. L. *Language and Psychology: Historical Aspect of Psycholinguistics*. New York: John Wiley. 1970. Pp. 248. £3·75.

BURROUGHS, G. E. R. *Design and Analysis in Educational Research*. University of Birmingham. 1971. Pp. 327. £2·00.

CHALMERS, N., CRAWLEY, R. AND ROSE, S. P. R. (Eds). *The Biological Bases of Behaviour*. London: Harper & Row Ltd. 1971. Pp. 318. £5·00.

DIXON, N. F. *Subliminal Perception: the Nature of a Controversy*. London: McGraw-Hill. 1971. Pp. 363. £4·00.

DOOB, L. W. *Patterning of Time*. Yale University Press. 1971. Pp. 472. £3·95.

ELKIND, D. *Children and Adolescents: Interpretive Essays on Jean Piaget*. Oxford University Press. 1971. Pp. 160. £2·70.

FINDLER, N. V. AND MELTZER, B. *Artificial Intelligence and Neuristic Programming*. Edinburgh University Press. 1971. Pp. 327. £6·00.

FJERDINGSTAD, E. J. (Ed.). *Chemical Transfer of Learned Information*. Amsterdam, London: North-Holland Publishing. 1971. Pp. 268. £7·91.

FOUCAULT, M. *Madness and Civilization: a History of Insanity in the Age of Reason*. London: Tavistock Publications. 1971. Pp. 299. £1·00.

GAZZANIGA, M. S. *The Bisected Brain*. New York: Appleton-Century-Crofts. 1970. Pp. 172. $12.10.

GLASER, R. (Ed.). *The Nature of Reinforcement*. New York: Academic Press. 1971. Pp. 379. £6·05.

GRUSSER, O.-J. AND KLINKE, R. (Eds). *Pattern Recognition in Biological and Technical Systems*. Berlin: Springer-Verlag. 1971. Pp. 413. $25.70.

HERTZ, R. H. *Frequency Tables for Scoring Rorschach Responses*. The Press of Case Western Reserve University: Cleveland & London. 1970. Pp. 274. $6.95.

HOLLANDER, E. P. *Principles and Methods of Social Psychology, 2nd edition*. Oxford University Press. 1971. Pp. 711. £4·75.

HUXLEY, R. AND INGRAM, E. *Language Acquisition: Models and Methods*. New York, London: Academic Press. 1971. Pp. 311. £5·00.

JARRARD, L. E. (Ed.). *Cognitive Processes of Nonhuman Primates*. New York, London: Academic Press. 1971. Pp. 188. £4·45.

JOHNSON, R. E. *Existential Man: the Challenge of Psychotherapy*. Oxford: Pergamon. 1971. Pp. 123. £2·25.

JONES, M., MCPHERSON, F., WHITAKER, D. S., SUTHERLAND, J. D., WALTON, H. AND WOLFF, H. *Small Group Psychotherapy.* Harmondsworth: Penguin Books Ltd. 1971. Pp. 128. £0·35.

JOURARD, S. M. *Self-disclosure : an Experimental Analysis of the Transparent Self.* New York: John Wiley. 1971. Pp. 248. £4·75.

KAGAN, J. *Understanding Children : Behavior, Motives and Thought.* New York: Harcourt Brace Jovanovich. 1971. Pp. 153. £1·25.

KINTSCH, W. *Learning, Memory, and Conceptual Processes.* New York, Chichester: John Wiley. 1970. Pp. 498. £4·40.

KLINGER, E. *Structure and Functions of Fantasy.* New York: Wiley. 1971. Pp. 424. £6·30.

KOHUT, H. *The Analysis of the Self : a Systematic Approach to the Psychoanalytic Treatment of Narcissistic Personality Disorders.* London: Hogarth Press. 1971. Pp. 368. £4·50.

LEE, W. *Decision Theory and Human Behavior.* New York: Wiley. 1971. Pp. 352. £6·15.

LEHRMAN, D. S., HINDE, R. A. AND SHAW, E. (Eds). *Advances in the Study of Behavior.* London: Academic Press. 1971. Pp. 263. £5·60.

LIEBERMAN, B. (Ed.). *Social Choice.* London, New York: Gordon & Breach. 1971. Pp. 425. £8·10.

LYNN, R. *Personality and National Character.* Oxford: Pergamon Press Ltd. 1971. Pp. 200. £3·75.

MACKINTOSH, N. J. AND HONIG, W. K. (Eds). *Fundamental Issues in Associative Learning.* Halifax: Dalhousie University Press. 1969. Pp. 203. $8.00.

MARTINI, L. AND GANONG, W. F. (Eds). *Frontiers in Neuroendocrinology.* 1971. Oxford University Press. 1971. Pp. 419. $25.00.

MORRIS, P. R. AND WHITING, H. T. A. *Motor Impairment and Compensatory Education.* London: G. Bell & Sons. 1971. Pp. 256. £2·50.

MROSOVSKY, N. *Hibernation and the Hypothalamus.* New York: Appleton-Century-Crofts. 1971. Pp. 287. $16.00.

NEEL, A. F. *Theories of Psychology.* London: University of London Press. 1971. Pp. 482. £4·75.

PEPITONE, A. *Attraction and Hostility : an Experimental Analysis of Interpersonal and Self Evaluation.* London: Tavistock Publications. 1971. Pp. 236. £0·70.

REICHARDT, W. (Ed.). *Processing of Optical Data by Organisms and Machines.* London: Academic Press. 1970. Pp. 614. £13·30.

RUBIN, F. *Learning and Sleep : the Theory and Practice of Hypnopaedia.* Bristol: John Wright & Sons. 1971. Pp. 138. £1·75.

SARTRE, J.-P. *Sketch for a Theory of the Emotions.* London: Methuen. 1971. Pp. 94. £0·50.

SCHMITT, F. O. (Ed.). *The Neurosciences : Second Study Program.* Rockefeller University Press. 1971. Pp. 961. $30.00.

SCHOENINGER, D. W. AND INSKO, C. A. *Introductory Statistics for the Behavioural Sciences.* Boston: Allyn & Bacon. 1971. Pp. 278. $8·95.

SEGAL, S. J. *Imagery : Current Cognitive Approaches.* New York, London: Academic Press Inc. 1971. Pp. 137. £3·50.

SLOBIN, D. I. (Ed.). *The Ontogenesis of Grammar.* New York, London: Academic Press. 1971. Pp. 247. £5·85.

SLUCKIN, W. (Ed.). *Early Learning and Early Experience.* Harmondsworth: Penguin Books Ltd. 1971. Pp. 416. £0·70.

STELLAR, E. AND SPRAGUE, J. M. (Eds). *Progress in Physiological Psychology, Vol.* 3. London: Academic Press. 1970. Pp. 273. £6·05.

SUMMERFIELD, A. (Scientific Editor). *Cognitive Psychology: British Medical Bulletin, Vol. 27, No. 3.* The British Council. 1971. Pp. 290. £2·00.

SUTHERLAND, N. S. AND MACKINTOSH, N. J. *Mechanisms of Animal Discrimination Learning.* New York, London: Academic Press. 1971. Pp. 559. £8·65.

TART, C. T. *On Being Stoned: a Psychological Study of Marijuana Intoxication.* Palo Alto, California: Science and Behavior Books. 1971. Pp. 333. $7.95.

VERNON, M. D. *Reading and its Difficulties.* Cambridge University Press. 1971. Pp. 211. £3·20.

WALCHER, D. N. AND PETERS, D. L. (Eds). *Early Childhood: the Development of Self-regulatory Mechanisms.* New York: Academic Press. 1971. Pp. 249. $11.00.

WARR, P. B. *Psychology at Work.* Harmondsworth: Penguin Books Ltd. 1971. Pp. 460. £0·75.

WOODWARD, W. M. *The Development of Behaviour.* Harmondsworth: Penguin. 1971. Pp. 208. £0·50.

PROCEEDINGS OF THE EXPERIMENTAL PSYCHOLOGY SOCIETY, 1971

29-30 March 1971. Meeting at Reading.

1st *Session*: "Enriched and impoverished rearing and the development of tactile discrimination in rats" by Dr D. Daniels* and Mr P. Mawer.* "Visual field and fixation changes after frontal eye-field lesions in monkeys" by Dr R. Latto and Dr A. Cowey. "Where in the brain does sodium amytal affect resistance to extinction?" by Dr J. Gray. "Behavioural suppression and functional brain nor-epinephrin" by Dr D. M. Warburton.

2nd *Session*: "Excursions in a motivational state plane" by Dr D. J. McFarland. "Studies of visual search processes (2) The effect of non-target heterogeneity density and distribution" by Mr G. Bloomfield. "Input labelling—a limitation to dividing attention between spatially separate visual inputs" by Jesus Alegria. "Parallel processing of single multi-dimensional stimuli" by Esther Saraga and Dr T. Shallice. "A new hypothesis on the nature of attention" by Dr D. A. Allport.

3rd *Session*: "Response mechanisms in transcription" by Dr H. J. Shaffer. "Repetition, contiguity and rehearsal in immediate memory" by Mr D. Routh. "Short-term memory and schizophrenia" by Helen Oddy and Professor P. H. Venables. "Recognition—an experimental study" by Professor G. Brown. "Responses of neurones in the cat's brain" by Miss A. C. Webb.*

4th *Session*: "What constitutes a speech sound" by Dr J. Morton and Susan Chambers.* "The detection of change in complex patterns in relation to antagonistic inhibition in the lateral geniculate body" by Mr W. A. Phillips* and Mr W. Singel.* "Discrimination of mirror images* by young children" by Dr P. E. Bryant. "A model for brightness contrast and depression" by Professor M. Treisman. "Reversed phi-movement" by Dr S. M. Anstis. "Visual dominance" by Mr H. F. Priest* and Mr A. A. Donnell.*

8-9 *July*, 1971. Meeting at Cambridge.

1st *Session*: "Visual and haptic recognition by children" by Dr J. Miller.* "Speech perception and psychological explanations" by Dr M. P. Haggard. "Visual localization in focal brain injury" by Dr G. Ratcliff* and Dr G. A. B. Davies-Jones.* "The cortico-retinal potential and alpha rhythm in the human electroencephalogram" by Dr O. C. J. Lippold* and Dr G. E. K. Novotny.*

2nd *Session*: "The perception of changes in stimulus wavelength: dynamic features of colour vision" by Dr D. Regan and Dr C. W. Tyler.* "The neural basis of the tilt after effect" by Dr M. Coltheart.* "Hemispheric differences in letter classification" by Dr G. Cohen.* "Recognition and retrieval of organized information" by Dr J. P. McLaughlin and Dr D. H. Herrman.* "High-speed scanning for words in primary and secondary memory" by Dr F. I. M. Craik.

3rd *Session*: "Hypothermia and behaviour of immunosympathectomized mice" by Dr C.

* by invitation

Van-Toller and Dr R. M. Tarpy.* "Observational learning in the laboratory mouse" by Mr M. R. Walley.* "Biochemical correlates of behaviour: a genetic approach" by Dr J. T. Rick. ". . . numbers languishing" by Dr M. Hammerton. "What is repeated in the repetition effect?" by Dr P. M. A. Rabbitt.

4th Session: Symposium on Skill: Chairman Dr E. C. Poulton. "Some studies in lateral asymetry" by Dr J. Annett and Mrs M. Annett.* "Timing the transfer of information between hemispheres in man" by Professor R. Davis. "The case of the virtuoso typist" by Dr L. H. Shaffer. "Relationship between speed and accuracy of paced movements" by Professor C. I. Howarth and W. D. A. Beggs.* Discussant: Dr D. E. Broadbent.

The Third Sir Frederick Bartlett Lecture:
"*Remembering* Revisited" by Professor O. L. Zangwill.

6–7 *January* 1972. Meeting at Birkbeck College, London.

1st Session: "Contrast and size variables and the tilt after-effect" by Mr D. M. Parker.* "Effective contrast constancy and apparent contrast" by Dr J. J. Kulikowski.* "Rods cancel cones in flicker" by Mr D. I. A. MacLeod.* "Temporal characteristics of stereoscopic vision: two eyes less sensitive than one" by Dr C. W. Tyler.* "When two ears are much better than one" by Dr G. B. Henning.*

2nd Session: "Obstacle detection with the aid of a direction noise generator" by Mr N. V. Clarke,* Mr G. F. Pick* and Dr J. P. Wilson. "Affective preference and mis-classification in a novel/familiar categorisation task" by Mr N. M. Freeman* and Mr D. M. Parker.* "Matching across touch and vision by human infants" by Dr P. E. Bryant, Miss P. Jones* and Miss V. Claxton.* "The spread of responses to 'meaningful' patterns in the cat's brain" by Miss A. C. Webb.* "Responses of neurones in the cat's visual cortex to relative movement of patterns" by Dr B. Delisle-Burns, Dr U. Gassinov* and Miss A. C. Webb.* "Is Helen an empiricist?" by Dr N. K. Humphrey.

3rd Session: "The serial position effect" by Mr G. Underwood.* "Attention strategies in memory" by Dr G. R. J. Hockey. "A new technique for examining organization in recall" by Mr A. Monk.* "Lexical search and phonemic organization in memory" by Dr P. T. Smith and Mr G. L. Claxton.*

4th Session: "Spatial and temporal tagging" by Dr N. O'Connor and Dr B. Hermelin. "The suppression of imagery" by Dr S. Jones. "Noise and the simultaneity of simultaneous interpretation" by Dr D. Gerver.* "New light on an old concept: vocabulary dynamics" by Dr D. H. Howes. "Tracking with a cross-coupled control system" by Mr P. D. McLeod.* "The interaction of heat with loss of sleep: task variables" by Dr E. C. Poulton.

Committee 1972
   President: Dr N. O'Connor.
   Editor: Professor R. Davis.
Ordinary Members of the Committee: Dr P. E. Bryant, Dr A. W. H. Buffery, Dr M. Hammerton, Dr D. Legge (Hon. Secretary), Dr J. Morton, Dr P. M. A. Rabbitt (Hon. Assistant Secretary), Dr L. H. Shaffer (Hon. Treasurer), Dr T. Shallice, Dr P. C. Watson.

* by invitation

The following have been elected to Ordinary Membership of the Society:

Dr J. R. Bloomfield, Department of Psychology, The University, Nottingham, NG7 2RD.

Dr O. J. Braddick, Psychological Laboratory, Downing Street, Cambridge, CB2 3EB.

Dr Gillian M. Cohen, Department of Experimental Psychology, South Parks Road, Oxford, OX1 3PS.

Dr T. J. Crow, Department of Mental Health, Medical School, Foresterhill, Aberdeen, AB9 2ZD.

Dr C. J. Darwin, Laboratory of Experimental Psychology, University of Sussex, Falmer, Brighton, Sussex, BN1 9QY.

Dr J. P. Frisby, Department of Psychology, The University, Sheffield, S10 2TN.

Dr D. N. Lee, Department of Psychology, The University, 60 The Pleasance, Edinburgh EH8 9TJ.

Dr Susanna Millar, Department of Experimental Psychology, South Parks Road, Oxford, OX1 3PS.

Dr A. D. Milner, Psychological Laboratories, The University, St. Andrews, Fife, KY16 9BA.

Dr J. D. Mollon, Department of Experimental Psychology, South Parks Road, Oxford, OX1 3PS.

Dr M. T. Morgan, Psychological Laboratory, Downing Street, Cambridge, CB2 3EB.

Dr R. E. Passingham, Department of Experimental Psychology, South Parks Road, Oxford, OX1 3PS.

Dr W. A. Phillips, Department of Psychology, The University, Stirling, Scotland.

Dr G. G. Ratcliff, Neuropsychology Unit, Department of Neurology, Churchill Hospital, Oxford, OX3 7IJ.

Dr A. J. Sanford, Department of Psychology, The University, Dundee, DD1 4HN.

Dr P. H. K. Seymour, Department of Psychology, The University, Dundee, DD1 4HN.

Dr F. M. Toates, Laboratory of Experimental Psychology, University of Sussex, Falmer, Brighton, Sussex, BN1 9QY.

Dr C. B. Trevarthen, Department of Psychology, The University, 60 The Pleasance, Edinburgh, EH8 9TJ.

Visiting Foreign Members

Professor R. C. Anderson, Department of Education, Research Council's Computer Based Learning Project, University of Leeds, Leeds, LS2 9JT.

Professor R. H. Day, Department of Psychology, Washington-Singer Laboratories, The University, Exeter, EX4 4QG.

Professor D. G. Doehring, Psychological Laboratory, Downing Street, Cambridge, CB2 3EB.

Professor A. Gilinski, Physiological Laboratory, Downing Street, Cambridge, CB2 3EG.

Dr L. Katz, Laboratory of Experimental Psychology, University of Sussex, Falmer, Brighton, Sussex, BN1 9QY.

Professor G. Mandler, Department of Experimental Psychology, South Parks Road, Oxford, OX1 3PS.

Dr J. Mandler, Department of Experimental Psychology, South Parks Road, Oxford, OX1 3PS.

Professor G. Mangan, Department of Experimental Psychology, South Parks Road, Oxford, OX1 3PS.

Dr J. McLaughlan, Department of Psychology, The University, 8–10 Berkeley Square, Bristol, BS8 1HH.

Professor P. Milner, Psychological Laboratory, Downing Street, Cambridge, CB2 3EB.

Professor C. Pribram, Department of Psychology, University College, Gower Street, London, WC1E 6BT.

Professor L. Riggs, Physiological Laboratory, Downing Street, Cambridge, CB2 3EG.

Professor J. R. Simon, Department of Psychology, The University, Sheffield, S10 2TN.

Professor Moncrief Smith, Psychological Laboratory, Downing Street, Cambridge, CB2 3EB.

Professor H-L. Teuber, Department of Experimental Psychology, South Parks Road, Oxford, OX1 3PS.

MEMBERS OF THE EXPERIMENTAL PSYCHOLOGY SOCIETY

Honorary Members

G. C. GRINDLEY, M.A., Psychological Laboratory, Downing Street, Cambridge, CB2 3EB.

W. E. HICK, M.B., 66 Bishop's Road, Cambridge, CB2 2NH.

Professor M. D. VERNON, M. A., Sc.D., 50 Cressingham Road, Reading, Berks, RG2 7TR

Ordinary Members

D. A. ALLPORT, Ph. D., Department of Psychology, The University, Earley Gate, White-knights, Reading, Berks, RG6 2AL.

D. P. ANDREWS, Ph.D., Department of Communications, The University, Keele, Staffs, ST5 5BG.

J. ANNETT, D.Phil., Department of Psychology, The University, Hull, HU6 7RX.

S. M. ANSTIS, Ph.D., Department of Psychology, The University, 8–10 Berkeley Square, BS8 1HH.

Professor R. J. AUDLEY, Ph.D., Department of Psychology, University College London, Gower Street, London, WC1E 6BT.

B. BABINGTON SMITH, O.B.E., M.A., Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford, OX1 3PS.

A. D. BADDELEY, Ph.D., Laboratory of Experimental Psychology, University of Sussex, Falmer, Brighton, BN1 9QY.

Professor P. BAKAN, Ph.D., Department of Psychology, Michigan State University, East Lansing, Michigan.

Professor H. B. BARLOW, M.D., M.B., B.Chir., F.R.S., Department of Physiology, 2570 Life Science Building, The University of California, Berkeley, California 94720, U.S.A.

Mrs E. BELBIN, Ph.D., Industrial Training Research Unit, 32 Trumpington Street, Cambridge, CB2 1QY.

A. J. BENSON, M.B., Ch.B., R.A.F. Institute of Aviation Medicine, Farnborough, Hants.

Professor D. E. BERLYNE, Ph.D., Department of Psychology, University of Toronto, Toronto 5, Ontario, Canada.

Professor P. BERTELSON, Ph.D., Laboratoire de Psychologie, Universitie Libre de Bruxelles, 115 Avenue Adolphe Buyl, Bruxelles, Belgium.

D. E. BLACKMAN, Ph.D., Department of Psychology, The University, P.O. Box 363, Birmingham, B15 2TT.

C. B. BLAKEMORE, B.Sc., Department of Psychology, Cane Hill Hospital, Coulsden, Surrey, CR3 3YL.

Colin B. BLAKEMORE, Ph.D., Physiological Laboratory, Downing Street, Cambridge, CB2 3EG.

J. R. BLOOMFIELD, Ph.D., Department of Psychology, The University, Nottingham, NG7 2RD.

T. M. BLOOMFIELD, Ph.D., Department of Psychology, York University, 4700 Keele Street, Downsview 463, Ontario, Canada.

R. A. BOAKES, Ph.D., Laboratory of Experimental Psychology, University of Sussex, Falmer, Brighton, BN1 9QY.

D. A. BOOTH, Ph.D., Laboratory of Experimental Psychology of Sussex, Falmer, Brighton, BN1 9QY.

R. BORGER, M.A., B.Sc., Department of Psychology, Brunel University, Uxbridge, Middlesex.

O. J. BRADDICK, Ph.D., Psychological Laboratory, Downing Street, Cambridge, CB2 3EB.

Professor G. S. BRINDLEY, M.D., F.R.S., Department of Physiology, Institute of Psychiatry, De Crespigny Park, London, SE5 8AF.

D. E. BROADBENT, Sc.D., F.R.S., MRC Applied Psychology Research Unit, 15 Chaucer Road, Cambridge, CB2 2EF.

Professor P. L. BROADHURST, Ph.D., Department of Psychology, The University, P.O. Box 363, Birmingham, B15 2TT.

I. D. BROWN, Ph.D., MRC Applied Psychology Research Unit, 15 Chaucer Road, Cambridge, CB2 2EF.

Professor J. BROWN, Ph.D., Department of Psychology, University of Bristol, 8–10 Berkeley Square, Bristol, BS8 1HH.

D. J. BRUCE, Ph.D., University Department of Education, 17 Brookside, Cambridge, CB2 1JG.

P. E. BRYANT, Ph.D., Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford, OX1 3PS.

A. W. H. BUFFERY, Ph.D., Human Development Research Unit, Park Hospital for Children, Old Road, Headington, Oxford.

F. W. CAMPBELL, M.B., Ch.B., Ph.D., D.O.M.S., Physiological Laboratory, Downing Street, Cambridge, CB2 3EG.

Miss V. R. CANE, M. A., Statistical Laboratory, Department of Mathematics, University of Manchester, M13 9PL.

A. CARPENTER, M.B., B.Chir., MRC Applied Psychology Research Unit, 15 Chaucer Road, Cambridge, CB2 2EF.

Professor E. C. CHERRY, D. Sc., M.I.E.E., Department of Electrical Engineering, Imperial College, Exhibition Road, London, SW7 2BT.

Gillian M. COHEN, D. Phil., Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford, OX1 3PS.

W. P. COLQUHOUN, Ph.D., MRC Applied Psychology Unit, Experimental Psychology Laboratory, University of Sussex, Falmer, Brighton, BN1 9QY.

Professor K. J. CONNOLLY, B.Sc., Department of Psychology, The University, Sheffield, S10 2TN.

R. CONRAD, Ph.D., MRC Applied Psychology Research Unit, 15 Chaucer Road, Cambridge, CB2 2EF.

D. W. J. CORCORAN, Ph.D., Department of Psychology, Faculty of Arts and Science, Memorial University of Newfoundland, St. Johns, Newfoundland, Canada.

A. COWEY, Ph.D., Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford, OX1 3PS.

J. J. COWLEY, Ph.D., Department of Psychology, Queen's University of Belfast, 1 Lennox-vale, Belfast, BT7 1NN.

F. I. M. CRAIK, Ph.D., Department of Psychology, Erindale College, University of Toronto, 3359 Mississauga Road, Clarkson, Ontario, Canada.

B. CRASKE, Ph.D., Department of Psychology, The University, Southampton, SO9 5NH.

Miss A. CRAWFORD, Ph.D., Department of Psychology, The University, P.O. Box 147, 7 Abercrombie Square, Liverpool, L69 3BX.

Professor E. R. F. W. CROSSMAN, Ph.D., Department of Industrial Engineering, University of California, Berkeley, California 94720, U.S.A.

T. J. CROW, Ph.D., Department of Mental Health, Medical School, Foresterhill, Aberdeen, AB9 2ZD.

S. CROWN, Ph.D., 5 Northwick Avenue, Harrow, Middx. HA3 OAA.

J. M. CULLEN, Ph.D., Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford, OX1 3PS.

H. C. A. DALE, Ph.D., Department of Psychology, The University, Hull, HU6 7RX.

C. J. DARWIN, Ph.D., Laboratory of Experimental Psychology, University of Sussex, Falmer, Brighton, Sussex, BN1 9QY.

D. R. Davies, Ph.D., Department of Psychology, The University, Leicester, LE1 7RH.

Professor R. Davis, D.Phil., Department of Psychology, The University, Claremont Place, Newcastle-upon-Tyne, NE1 7RU.

B. Delisle Burns, M.R.C.S., L.R.C.P., Division of Physiology, National Institute for Medical Research, Mill Hill, London, NW7 1AA.

J. D. Delius, Ph.D., Department of Psychology, University of Durham, Science Laboratories, South Road, Durham.

Professor J. A. Deutsch, D.Phil., Department of Psychology, University of California, San Diego, La Jolla, California 92037, U.S.A.

S. J. Dimond, Ph.D., Department of Psychology, University College, Cardiff, CF1 1XL.

N. F. Dixon, Ph.D., Department of Psychology, University College London, Gower Street, London, WC1E 6BT.

Professor P. C. Dodwell, Ph.D., Department of Psychology, Queen's University, Kingston, Ontario, Canada.

J. Drever, M. A., Principal, University of Dundee, College Gate, Dundee, DD1 4HN.

Professor G. C. Drew, M.A., Department of Psychology, University College London, Gower Street, London, WC1E 6BT.

Professor J. T. Eayrs, D.Sc., Department of Anatomy, The Medical School, Birmingham, B15 2TJ.

A. Elithorn, M.D., M.B., B.Chir., M.R.C.P., D.P.M., Hut J, Royal Free Hospital, Lawn Road, London, NW3 2XJ.

G. Ettlinger, Ph.D., Institute of Psychiatry, De Crespigny Park, Denmark Hill, London, SE5 8AF.

J. M. Findlay, Ph.D., Department of Psychology, University of Durham, Science Laboratories, South Road, Durham.

Professor B. M. Foss, M.A., Department of Psychology, Bedford College, Regent's Park, London, NW1 4NS.

Professor E. Fraser, Ph.D., Department of Psychology, Taylor Building, The University, Aberdeen, AB9 2UB.

J. P. Frisby, Ph.D., Department of Psychology, The University, Sheffield, S10 2TN.

J. L. Gedye, M.B., B.Chir., Department of Electrical Engineering Science, University of Essex, Colchester, CO4 3SQ.

F. H. George, Ph.D., Department of Cybernetics, Brunel University, Uxbridge, Middlesex.

Mrs F. Goldman-Eisler, D.Phil., Psycholinguistics Research Unit, Department of Phonetics, University College London, Gower Street, London, WC1E 6BT.

I. E. Gordon, Ph.D., Department of Psychology, Washington-Singer Laboratories, Exeter, EX4 4QG.

G. W. Granger, Ph.D., Department of Psychology, University of Durham, Science Laboratories, South Road, Durham.

J. A. Gray, Ph.D., Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford, OX1 3PS.

Judith M. Greene, Ph.D., Department of Psychology, Birkbeck College, Malet Street, London, WC1E 7HX.

A. H. Gregory, Ph.D., Department of Psychology, The University, Manchester, M13 9PL.

Professor R. L. Gregory, M.A., Brain and Perception Laboratory, Department of Anatomy, Medical School, University Walk, Bristol, BS8 1TD.

Professor R. A. M. Gregson, Ph.D., Department of Psychology and Sociology, University of Canterbury, P.O. Box 1471, Christchurch, C.I., New Zealand.

Professor S. Griew, Ph.D., Department of Psychology, The University, Dundee, DD1 4HN.

Professor M. P. Haggard, Ph.D., Department of Psychology, Queen's University of Belfast, 1 Lennoxvale, Belfast, BT7 1NN.

D. Hale, M.Sc., Department of Psychology, Queen's University of Belfast, 1 Lennoxvale, Belfast, BT7 1NN.

A. M. HALLIDAY, M.B., Ch.B., M.R.C. Neurological Unit, National Hospital, Queen Square, London, WC1N 3BG.

M. S. HALLIDAY, Ph.D., Laboratory of Experimental Psychology, University of Sussex, Falmer, Brighton, Sussex, BN1 9QY.

M. HAMMERTON, Ph.D., MRC Applied Psychology Research Unit, 15 Chaucer Road, Cambridge, CB2 3EB.

R. HARPER, Ph.D., Department of Food Science, University of Reading, Reading, RG1 5QA.

A. D. HARRIS, M.R.C.S., L.R.C.P., MRC Psych., D.P.M., Littlemore Hospital, Oxford, OX4 4XN.

G. J. HARRISON, Ph. D., Department of Psychology, University College, P.O. Box 78, Cardiff, CF1 1XL.

Miss A. W. HEIM, Ph.D., Psychological Laboratory, Downing Street, Cambridge, CB2 3EB.

L. J. HERBERG, M.B., B.Ch., D.Phil., Institute of Neurology, National Hospital, Queen Square, London, WC1N 3BG.

Mrs B. M. HERMELIN, Ph.D., MRC Developmental Psychology Research Unit, Drayton House, Gordon Street, WC1H OAN.

R. HINCHCLIFFE, M.D., Institute of Laryngology, Royal National Hospital, Gray's Inn Road, London, WC1X 8DA.

Professor R. A. HINDE, D.Phil., MRC Unit on the Development and Integration of Behaviour, High Street, Madingley, Cambridge, CB3 8AA.

G. R. J. HOCKEY, Ph.D., Department of Psychology, Science Laboratories, South Road, Durham.

Professor D. H. HOLDING, Ph.D., Department of Psychology, University of Louisville, Kentucky 40208, U.S.A.

C. M. HOLLOWAY, Ph.D., MRC Applied Psychology Research Unit, 15 Chaucer Road, Cambridge, CB2 2EF.

G. HORN, M.B., Ch.B., University of Cambridge, Department of Anatomy, Cambridge, CB2 3DY.

N. HOTOPF, M.A., Department of Social Psychology, London School of Economics, Houghton Street, Aldwych, London, WC2A 2AE.

Professor I. P. HOWARD, Ph.D., Department of Psychology, York University, Keele Street, Toronto 12, Ontario, Canada.

Professor C. I. HOWARTH, D.Phil., Department of Psychology, The University, Nottingham, NG7 2RD.

J. A. M. HOWE, Ph.D., Bionics Research Laboratory, School of Artificial Intelligence, University of Edinburgh, Forrest Hill, Edinburgh, EH1 2QL.

D. H. HOWES, Ph.D., Room E6-121 (Neurology), Boston Veterans Administrative Hospital, 150 South Huntington Avenue, Boston Massachusetts 02130 U.S.A.

I. M. HUGHES, Ph.D., Department of Psychology, Taylor Building, The University, Aberdeen, AB9 2UB.

N. K. HUMPHREY, Ph.D., Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford, OX1 3PS.

Professor I. M. L. HUNTER, D. Phil., Department of Psychology, The University, Keele, Staffs, ST5 5BG.

Professor H. M. B. HURWITZ, Ph.D., Department of Psychology, College of Social Science, University of Guelph, Guelph, Ontario, Canada.

Miss M. IKEDA, Ph.D., Research Department of Ophthalmology, Royal College of Surgeons, Lincoln's Inn Fields, London, WC2A 3PM.

J. G. INGHAM, Ph.D., Medical Research Council External Staff, Llandough Hospital, Penarth, Glam., CF6 1XW.

J. INGLIS, Ph.D., Department of Psychology, Queen's University, Kingston, Ontario, Canada.

Mrs S. D. IVERSEN, Ph.D., Psychological Laboratory, Downing Street, Cambridge, CB2 3EB.

Professor P. H. R. JAMES, Ph.D., Department of Psychology, Dalhousie University, Halifax, Nova Scotia, Canada.

Professor M. A. JEEVES, Ph.D., Psychology Department, University of St. Andrews, St. Andrews, Fife, Scotland.

P. N. JOHNSON-LAIRD, Ph.D., Department of Psychology, University College London, Gower Street, London, WC1E 6BT.

A. R. JONCKHEERE, Ph.D., Department of Psychology, University College London, Gower Street, London, WC1E 6BT.

D. JONES, Ph.D., Department of Psychology, Birkbeck College, Malet Street, London, WC1E 7HX.

Mrs S. JONES, Ph.D., Department of Psychology, University College London, Gower Street, London, WC1E 6BT.

Mrs D. JOYCE, Ph.D., Department of Psychology, Birkbeck College, Malet Street, London, WC1E 7HX.

Professor H. KAY, Ph.D., Department of Psychology, The University, Sheffield, S10 2TN.

R. P. KELVIN, Ph.D., Department of Psychology, University College London, Gower Street, London, WC1E 6BT.

C. C. KIERNAN, Ph.D., Department of Child Development, Institute of Education, Malet Street, London, WC1E 7HS.

M. KINSBOURNE, B.M., B.Ch., M.R.C.P., Department of Pediatrics, Box 2948, Duke Medical Centre, Durham, N.C., 27706, U.S.A.

G. R. KISS, B.Sc., MRC Speech and Communication Research Unit, University of Edinburgh, 31 Buccleuch Place, Edinburgh, EH8 9JT.

D. R. J. LAMING, Ph.D., Psychological Laboratory, Downing Street, Cambridge, CB2 3EB.

R. LATTO, Ph.D., Psychological Laboratory, Downing Street, Cambridge, CB2 3EB.

D. N. LEE, Ph.D., Department of Psychology, The University, 60 The Pleasance, Edinburgh, EH8 9TJ.

D. LEGGE, Ph.D., Department of Psychology, Birkbeck College, Malet Street, London, WC1E 7HX.

Professor B. N. LEWIS, B.Sc., Institute of Educational Technology, The Open University, Bletchley, Bucks.

S. A. LEWIS, Ph.D., University Department of Psychiatry, Royal Edinburgh Hospital, Morningside Park, Edinburgh, EH10 5HF.

W. A. LISHMAN, B.Sc., M.D., M.R.C.S., M.R.C.P., A.P.M., Maudsley Hospital, Denmark Hill, London, SE5 8AZ.

N. E. LOVELESS, Ph.D., Department of Psychology, The University, Dundee, DD1 4HN.

D. J. McFARLAND, D.Phil., Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford, OX1 3PS.

Professor D. M. MACKAY, Ph.D., Department of Communications, The University, Keele, Staffs., ST5 5BG.

Professor N. J. S. M. MACKINTOSH, Ph.D., Department of Psychology, Dalhousie University, Halifax, Nova Scotia, Canada.

N. H. MACKWORTH, M.B., Ch.B., Ph.D., 429 Atlantic Street, Stamford, Conn., U.S.A.

Miss F. E. MACNEILL, Ph.D., Department of Psychology, The University, Oxford Road, Manchester, M13 9PL.

E. M. MACPHAIL, D.Phil., Laboratory of Experimental Psychology, University of Sussex, Falmer, Brighton, Sussex, BN1 9QY.

J. C. MARSHALL, Ph.D., MRC Speech and Communication Research Unit, The University, 31 Buccleuch Place, Edinburgh, EH8 9JT.

Mrs I. MARTIN, Ph.D., Institute of Psychiatry, De Crespigny Park, Denmark Hill, London, SE5 8AF.

W. A. MATTHEWS, B.A., Department of Psychology, The University, Southampton, SO9 5HN.

Miss S. MAYNE, Ph.D., 88 Bulmershe Road, Reading, Berks, RG1 5RJ.

Professor K. A. PROVINS, Ph.D., Department of Psychology, Australian National University, Box 4, P.O. Canberra, A.C.T. 2600, Australia.

P. M. A. RABBITT, Ph.D., Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford, OX1 3PS.

G. G. RATCLIFF, D. Phil., Neuropsychology Unit, Department of Neurology, Churchill Hospital, Oxford, OX3 7IJ.

D. REGAN, Ph.D., Department of Communication, The University, Keele, Staffs., ST5 5BG.

Professor R. L. REID, M.A., Psychological Laboratory, The University, Washington-Singer Laboratories, Exeter, EX4 4QG.

J. T. RICK, Ph.D., Department of Psychology, University of Birmingham, P.O. Box 363, Birmingham, B15 2TT.

Miss H. E. ROSS, Ph.D., Department of Psychology, The University, Stirling, Scotland.

Professor W. A. H. RUSHTON, Ph.D., Sc.D., M.R.C.P., L.R.C.P., F.R.S., Trinity College, Cambridge, CB2 1TQ.

I. S. RUSSELL, Ph.D., Department of Psychology, University College London, Gower Street, London, WC1E 6BT.

Professor R. W. RUSSELL, Ph.D., D.Sc., University of California, Irvine, Calif., 92664, U.S.A.

Professor D. RUSSELL DAVIS, M.D., M.B., B.Chir., M.R.C.P., M.R.C.S., L.R.C.P., D.P.M., 39–41 St. Michael's Hill, Bristol, BS2 8DZ.

E. A. SALZEN, Ph.D., Department of Psychology, Taylor Building, The University, Aberdeen, AB9 2UB.

S. J. SANFORD, Ph.D., Department of Psychology, The University, Dundee, DD1 4HN.

B. SEMEONOFF, Ph.D., Department of Psychology, University of Edinburgh, 60 The Pleasance, Edinburgh, EH8 9TJ.

P. H. K. SEYMOUR, Ph.D., Department of Psychology, The University, Dundee, DD1 4HN.

L. H. SHAFFER, Ph.D., Psychological Laboratory, The University, Washington-Singer Laboratories, Exeter, EX4 4QG.

T. SHALLICE, Ph.D., Department of Psychology, University College London, Gower Street, London, WC1E 6BT.

M. H. SHELDON, Ph.D., Department of Psychology, The University, Earley Gate, Whiteknights, Reading, RG6 2AL.

M. E. SIME, B.A., Department of Psychology, The University, Sheffield, S10 2TN.

Professor W. SLUCKIN, Ph.D., Department of Psychology, University of Leicester, Leicester, LE1 7RH.

P. T. SMITH, D.Phil., Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford, OX1 3PS.

J. R. SMYTHIES, M.Sc., M.D., M.B., B.Chir., M.R.C.P., D.P.M., University Department of Psychiatry, Royal Edinburgh Hospital, Morningside Park, Edinburgh, EH10 5BU.

A. C. STANILAND, Ph.D., Department of Psychology, The University, Washington-Singer Laboratories, Exeter, EX4 4QG.

Professor H. STEINBERG, Ph.D., Department of Pharmacology, University College London, Gower Street, London, WC1E 6BT.

Miss J. G. STEVENSON, Ph.D., Psychological Laboratory, Downing Street, Cambridge, CB2 3EB.

R. G. A. STRETCH, Ph.D., Department of Psychology, University of Saskatchewan, Saskatoon, Saskatchewan, Canada.

A. W. STILL, Ph.D., Department of Psychology, Science Laboratories, The University, South Road, Durham.

H. C. W. STOCKBRIDGE, M.A., Quest, Tilford, Farnham, Surrey.

Professor A. SUMMERFIELD, B. Sc., Department of Psychology, Birkbeck College, Malet Street, London, WC1E 7HX.

Professor N. S. SUTHERLAND, D.Phil., Laboratory of Experimental Psychology, University of Sussex, Falmer, Brighton, BN1 9QY.

Professor J. R. SYMONS, Ph.D., Department of Psychology, King's College, Taylor Building, Aberdeen, AB9 2UB.

Professor J. SZAFRAN, Ph.D., Department of Psychology, University of Southern California, University Park, Los Angeles, California, 90007, U.S.A.

Professor H. TAJFEL, M.A., Department of Psychology, The University, 8–10 Berkeley Square, Bristol, BS8 1HH.

D. H. TAYLOR, Ph.D., Department of Psychology, The University, Southampton, SO9 5NH.

W. B. TEMPLETON, M.Sc., Ph.D., Department of Psychology York University, Keele Street, Toronto 12, Ontario, Canada.

E. A. C. THOMAS, Ph.D., Department of Psychology, University of Michigan, Ann Arbor, Michigan 48104, U.S.A.

Mrs B. TIZARD, Ph.D., Department of Child Development, Institute of Education, Malet Street, London, WC1E 7HS.

F. M. TOATES, D.Phil., Laboratory of Experimental Psychology, University of Sussex, Falmer, Brighton, Sussex, BN1 9QY.

C. B. TREVARTHEN, Ph.D., Department of Psychology, The University, 60 The Pleasance, Edinburgh, EH8 9TJ.

Mrs A. TREISMAN, D. Phil., Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford, OX1 3PS.

Professor M. TREISMAN, M.B., B.Ch., Department of Psychology, The University, Earley Gate, Whiteknights, Reading, RG6 2AL.

Professor E. L. TRIST, O.B.E., M.A., Graduate School of Business Administration (Management Area), University of California, Los Angeles, California 90024, U.S.A.

Professor A. M. UTTLEY, Ph.D., Laboratory of Experimental Psychology, University of Sussex, Falmer, Brighton, Sussex, BN1 9QY.

Professor P. H. VENABLES, Ph.D., Department of Psychology, Birkbeck College, Malet Street, London, WC1E 7HX.

Miss M. A. VINCE, M.A., Psychological Laboratory, Downing Street, Cambridge, CB2 3EB.

Professor D. VOWLES, D.Phil., Department of Psychology, The University, 60 The Pleasance, Edinburgh, EH8 9TJ.

G. K. WALLACE, Ph.D., Department of Psychology, The University, Earley Gate, Whiteknights, Reading, RG6 2AL.

D. M. WARBURTON, Ph.D., Department of Psychology, The University, Earley Gate, Whitenights, Reading, RG6 2AL.

P. B. WARR, Ph.D., Department of Psychology, The University, Sheffield, S10 2TN.

Mrs E. WARRINGTON, Ph.D., Psychological Department, National Hospital, Queen Square, London, WC1N 3BG.

P. C. WASON, Ph.D., Department of Phonetics, University College London, Gower Street, London, WC1E 6BT.

A. J. WATSON, M.A., Psychological Laboratory, Downing Street, Cambridge, CB2 3EB.

R. H. J. WATSON, Ph.D., Department of Nutrition, Queen Elizabeth College, Campden Hill Road, London, W8 7AH.

Miss K. P. WATTS, Psychological Laboratory, Downing Street, Cambridge, CB2 3EB.

Professor L. WEISKRANTZ, Ph.D., Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford, OX1 3PS.

Professor A. T. WELFORD, Sc.D., Department of Psychology, University of Adelaide, Adelaide, South Australia.

T. C. D. WHITESIDE, M.B.E., Ph.D., M.B., Ch.B., R.A.F. Institute of Aviation Medicine, Farnborough, Hants.

J. W. WHITFIELD, M.A., Department of Psychology, University College London, Gower Street, London, WC1E 6BT.

P. WHITTLE, Ph.D., Psychological Laboratory, Downing Street, Cambridge, CB2 3EB.

R. T. WILKINSON, Ph.D., MRC Applied Psychology Resarch Unit, 15 Chaucer Road, Cambridge, CB2 2EF.

The Hon. Mrs M. WILLIAMS, Ph.D., Department of Psychology, United Cambridge Hospitals, 2 Bene't Place, Cambridge, CB2 1EL.

J. A. WILSON, B.Sc., Department of Psychology, University College, P.O. Box 78, Cardiff, CF1 1XL.

J. P. WILSON, Ph.D., Department of Communication, The University, Keele, Staffs., ST5 5BG.

Mrs M. WYKE, Ph.D., Department of Psychological Medicine, Maida Vale Hospital, London, W9 1TL.

Professor O. L. ZANGWILL, M.A., Psychological Laboratory, Downing Street, Cambridge, CB2 3EB.

Methods in Psychobiology
Volume 1

Laboratory Techniques
in Neuropsychology and Neurobiology
edited by R. D. Myers

Students and researchers in psychology, physiology, and related disciplines have long felt the need for a unified collection of information presenting laboratory methods used in experiments which relate the brain to behaviour. *Methods in Psychobiology* aims to fill this gap. Volume 1 presents, in great detail, a wide variety of basic procedures. Experiments are described step by step and each author explains the practical pitfalls to avoid as well as the advantages and limitations of each technique from a theoretical standpoint. Photographs and diagrams are produced in many chapters to assist the beginner. In the chapter on histology, coloured photographs of histological sections are included that may be used as models. Each chapter is followed by an appendix of manufacturers and suppliers, so that specialized materials and equipment may easily be obtained.

This book is the first volume of what promises to be a highly informative work, and should not only be used by beginners in the field, but consulted by experienced scientists requiring basic techniques for their own line of research.

Contents

Principles of research in physiological psychology. General laboratory procedures. Use of the stereotaxic technique. Neurosurgical procedures with special reference to aspiration lesions. Introducing subcortical lesions by electrolytic methods. Recording changes in electrical properties in the brain: the EEG. Electrical stimulation of the brain. Methods for chemical stimulation of the brain. Elementary histology for neuropsychologists. Determining changes in vital functions: ingestion. Author index. Subject index.

In preparation:

Volume 2: Specialized Laboratory Techniques in Neuropsychology and Neurobiology.

Volume 3: Advanced Laboratory Techniques in Neuropsychology and Neurobiology.

Feedback Mechanisms in Animal Behaviour

D. J. McFarland

Balliol College, Oxford, England
December 1971, xii + 280 pp., £3·80

The work aims to introduce biologists who have a limited knowledge of mathematics to the principles of those aspects of control theory relevant to the study of animal behaviour. An attempt is made to facilitate an intuitive understanding of these principles without introducing higher mathematical concepts such as complex numbers. Simultaneously, advanced aspects of control theory, such as optimization theory, which are highly relevant to the study of behaviour are dealt with.

In applying control theory to aspects of behaviour not previously considered in this light (some issues such as motivational energy being reinterpreted from this point of view), and similarly illustrating control theory concepts with behavioural examples, the author covers important territory between these two disciplines. Aimed primarily at those working in the sphere of experimental psychology, this book will also be of great interest both to biologists concerned with control and physical scientists interested in behaviour.

Contents

An outline of control theory. Open and closed loop control. Behaviour as a function of time. Behaviour as a function of frequency. The accuracy of behavioural control. Black box analysis. Control theory and behaviour theory. References and author index. Subject index.

Microcirculation, Perfusion, and Transplantation of Organs

Edited by **THEODORE I. MALININ**
Georgetown University
School of Medicine-Dentistry
Department of Research Pathology
Washington, D.C.

BERNARD S. LINN
Department of Surgery
University of Miami School of Medicine and the
Veterans' Administration Hospital
Miami, Florida

ARTHUR B. CALLAHAN
Medicine and Dentistry Program
Office of Naval Research
Washington, D.C.

W. DEAN WARREN
Department of Surgery
University of Miami School of Medicine
Miami, Florida

This book presents critical and authoritative papers on current work being done in the area of microcirculation. The informative papers included in the book illustrate the relationship of microcirculation to organ transplantation and organ perfusion. Research and experimental work in each of these fields is tied together by the editors.

The initial section of this book is devoted to microcirculation under normal conditions and to pathology and pathophysiology of microcirculation. The microcirculation as the key to organ transplantation and perfusion is discussed in the last sections. Each section is concluded with summary papers and overviews by special discussants. This book will be a valuable link between basic scientists interested in microcirculation and surgeons concerned with organ transplantation and preservation, as well as a valuable basic reference text for these related fields.

1970, 423 pp., £5·85

CONTENTS

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY

Volume 24   Part 2   May 1972

PUBLISHED FOR THE EXPERIMENTAL PSYCHOLOGY SOCIETY BY ACADEMIC PRESS LONDON AND NEW YORK

REMEMBERING REVISITED†

O. L. ZANGWILL

Cambridge Psychological Laboratory

Introduction

This, inevitably, is a sad occasion. Although I am privileged to give the third Sir Frederic Bartlett Lecture it is, in effect, the first Sir Frederic Bartlett Memorial Lecture. As everyone here will know, Sir Frederic died in Cambridge after a short illness on 30 September, 1969, within three weeks of his 83rd birthday. He had begun work in the Cambridge Psychological Laboratory in 1914, becoming its Director in 1922 and the first Professor of Experimental Psychology in the University in 1931. Sir Frederic virtually created the Medical Research Council's Applied Psychology Unit, established in the Department of Experimental Psychology in 1944 and now our flourishing sister laboratory in Chaucer Road. After his retirement in 1952, Sir Frederic remained closely associated with the affairs of Psychology in Cambridge almost to the end.

I do not propose to speak of Sir Frederic's great contribution to the development of experimental psychology in Britain or of the work of the Department which he built up and fostered over a period of some 30 years. The magnitude of his contribution has been widely acclaimed, though perhaps something still remains to be said of his extraordinary personal influence. I have given myself the more limited aim of looking back on some of his earlier work, particularly that brought together in his famous book *Remembering*, published in 1932, and trying to trace something of the influence which it has had on the approach of members of my own generation to psychological problems. I may perhaps add that I was myself an undergraduate in his Department not long after the book appeared and remember vividly the lively discussions it provoked, in which—characteristically—the author was often his own most incisive critic. In revisiting *Remembering*, I hope to steer a course midway between the twin hazards of ancestor worship on the one hand and the war of the generations on the other. I was—and am—deeply impressed by *Remembering* but, if some of my remarks should strike a critical note, I do not think this would have been taken by my predecessor in office to imply disrespect. It was he, after all, who taught us how to criticize.

The Origins of Bartlett's Work

In his Preface to *Remembering*, Bartlett relates that his work had its beginnings in a demonstration he had been asked to give on the occasion of the formal opening of the Cambridge Psychological Laboratory in May, 1913. This, he tells us, took place on "a brilliant afternoon", though I may say that I was once told by the

†The third Sir Frederic Bartlett Lecture delivered in Cambridge on 9 July 1971

late Professor Sir Godfrey Thomson that, in fact, it was pouring with rain! What-ever the vicissitudes of the Cambridge climate, however, there can be no doubt that the demonstration *did* take place and that it was concerned with visual percep-tion under conditions of short exposure. "It was clear", wrote Bartlett, "that the course of visual perception may be determined by a very large number of different factors and . . . that a carefully arranged experimental study might serve to dis-entangle many of these factors and to demonstrate their functions."

Bartlett set to work at once. In his earliest experiments, geometrical forms, patterns and designs, some of them forming a progressive series, were exposed for a quarter of a second and the subjects asked to reproduce what they had seen by drawing, where necessary supplemented by description. Pictorial material of varying complexity was also used. These experiments were reported in a paper published in the *British Journal of Psychology* (Bartlett, 1916a) and, in briefer form, in an early chapter of *Remembering*. In both accounts, Bartlett analyses in some detail the results obtained, placing particular emphasis on the ways in which the subject actively organizes the material presented to him. This he describes, somewhat figuratively, as "effort after meaning"—an unwitting attempt to match whatever is presented to some pre-existing setting or scheme and which does not necessarily imply felt effort or strain on the part of the percipient. In his opinion the traditional view of perception as the passive intake of sensory information can no longer be sustained. On the contrary, perceiving is to be envisaged as an active, dynamic process, directed by interest and feeling, and inextricably linked with imaging, valuing and the beginnings of judgement.

Inevitably, this study of perception led Bartlett to a concern with remembering, of which, as he said, "perception may legitimately be regarded as the psychological starting point". Remembering he approached initially through the study of what he called *imaging*, i.e. the responses that may be evoked when one contemplates amorphous material, as in the old pastime of seeing faces in the fire. As his material Bartlett chose inkblots, later to be endowed with dubious celebrity by Hermann Rorschach, whose excursion into psychodiagnostics Bartlett always seemed to view with faint distaste. His own interest in imaging arose purely from the belief that it constitutes a kind of bridge between perceiving, in which direct sense stimulation is present, and remembering, in which it is not. Further, he was evidently fas-cinated by the diversity of human interests provoked by this kind of material and their apparent relation to individual differences in types of imagery. To this topic I shall return later.

This, then, is the background of Bartlett's work on memory. While, in the Preface to *Remembering*, he talks of having first, "as if in duty bound", followed the lead of Ebbinghaus and worked for some time with nonsense material, I am doubt-ful whether the "disappointment and growing dissatisfaction" of which he speaks had much to do with the methods of inquiry he finally chose. No one who had been a pupil of Ward and a disciple of Rivers could be expected to embrace the Ebbinghaus methods with any enthusiasm. From the beginning, Bartlett was evidently seeking methods more flexible, more life-like and more appropriate to bring out the characteristics of remembering envisaged as an ongoing individual and social activity.

Experiments on Memory

Bartlett's experiments on memory were carried out in the main between 1914 and 1916 and formed the core of his successful Fellowship Dissertation submitted to St John's College, Cambridge, in the latter year (Bartlett, 1916b). These, too, reappear in *Remembering* without great change of form, though with considerable revision and the addition of much new material. They are also given a radically new theoretical interpretation.

The methods themselves are very simple, perhaps deceptively so. Simplicity, as Bartlett (1932) himself reminded us, is an elusive concept, at any rate in psychology. As he pointed out, it is not necessarily achieved by the use of artificially simplified stimuli or by conditions which unduly constrain or limit the subject's response. In consequence, his own methods were, even for the time, somewhat informal, and by modern standards leave much to be desired. In the first, which he called the *Method of Description*, the subject was asked to describe a series of five drawings of faces which had been briefly shown to him and to answer a series of questions about them. This interrogation was repeated a week or two later. In the second, the *Method of Repeated Reproduction*, short stories varying in length between 200 and 500 words were read through twice by the subjects, who reproduced them in writing from memory after 15 min and again at varying intervals over a period of weeks or even months. Some of the stories were taken from Indian and African sources. In other experiments, pictures were used but the results of these are not presented in *Remembering*. Bartlett freely admits that he owes this method to the French psychologist, Jean Philippe, who in 1897 had published a paper on the transformation of mental images. In fact, however, Philippe used common objects, not stories or pictures, and a combination of tactual and visual presentation. He also worked principally with children. In my view, Bartlett's Method of Repeated Reproduction should be regarded not only as original but as the source of much that is most valuable in his work.

The two remaining methods were chosen with social factors more specifically in mind. The *Method of Picture Writing* was undoubtedly suggested to Bartlett by his interest in *conventionalization*, a term adapted from Rivers (1912) to denote the changes undergone by representational material when it passes from one culture to another. In Bartlett's experiment, the subjects were required to learn to substitute a set of pictorial signs for specified words, tests being given 15 min and two weeks after the end of the learning session. Finally, the *Method of Serial Reproduction* was evidently devised as an analogue to the transmission of information through a social group, as for example in the case of rumour. Material, verbal or pictorial, reproduced by one subject is presented for reproduction by a second, his reproduction presented to a third and so on until a chain of reproductions has been obtained. In *Remembering*, Bartlett devotes by far the greatest amount of space to the material obtained by this method, discussing different types of passage, narrative, descriptive and argumentative, and the results obtained with different racial groups. He clearly regarded it as having much to contribute to our understanding not only of the influence of cultural factors on memory but also of communication within and between widely different social groups.

What were the main findings from these experiments? These are hard to

summarize, largely because the results are highly selected and presented wholly in qualitative form. The most outstanding finding, however, is the extraordinarily high proportion of inaccuracy occurring in all reproductions, not excluding the first, and the fact that the subjects appeared to be totally unaware of the extent of their inaccuracy. These features appear most strikingly in the results of the experiments on the repeated and serial reproduction of stories, from which it becomes clear that inaccuracy results not only from omission and condensation, as might have been expected, but from radical transformation of the original material. Style is poorly conveyed, irrelevant material discarded, and what is retained undergoes marked, persistent and sometimes progressive rationalization. These changes are particularly apparent in the case of stories originating in alien cultures, in which the whole of the narrative may be recast in a form compatible with the subject's own cultural background and social conventions. They are also very striking in the experiments on serial reproduction of pictures. The various forms which transformation in memory may take were analysed by Bartlett with great perceptiveness and skill and some of the factors upon which they depend, both individual and social, discussed at considerable length. His main conclusion is that remembering can in no sense be regarded as the mere revival of earlier experience; it is a process of active reconstruction, much of it based on factors of general impression and attitude, together with the reinstatement of a small amount of critical detail. In his own words, recall is far more decisively an affair of construction than one of mere reproduction.

Bartlett's Theory of Memory

In his Fellowship Dissertation, Bartlett (1916b) devoted much discussion to the interrelation of factors leading to reproduction and to their interpretation in terms of the largely introspective psychology of the period. He was evidently impressed by Margaret Washburn's *Movement and Mental Imagery* (1916) and had a good deal to say about the views of such German psychologists as Müller-Freienfels (1913) and Betz (1910). Indeed Bartlett's use of the term *attitude* owes much to Betz's term *Einstellung*, which broadly signifies mental posture or set. He was impressed, too, by the Würzburg work on imageless thought (in which incidentally he appeared fully to believe) and had much of interest to say on the general topic of conventionalization. With the exception of the last, however, these topics do not reappear in *Remembering*, suggesting perhaps a weakening of the Continental—in particular German—influence and a correlated retreat from the more tortuous ramifications of introspectionism. But in this book there is great development: Bartlett's interest in social anthropology, always keen, had led him to take an increasing interest in social aspects of psychology, quickened perhaps by his visit to South Africa in 1929. It is not surprising, therefore, to find the second part of *Remembering* wholly devoted to memory in its social and cross-cultural aspects. Even more important, perhaps, is that he had come to know Henry Head and had had many wide-ranging discussions with him during the writing of Head's classical monograph on aphasia (Head, 1926). This relationship with Head appears to have given Bartlett the essential intellectual stimulus which he needed to construct a general theory of memory.

As is well known, Bartlett regarded his experimental findings as incompatible with any notion of what, following Philippe, he called "fixed, lifeless traces", that is to say, *simulacra* of past events in some way preserved in their original form and whose re-excitation constitutes recall. In consequence, he tried to replace the conventional "trace theory" of memory, which in one form or another has been with us since Plato (Gomulicki, 1953*b*), by a theory of *schemata*, adapted from the work of Head and Holmes (1911–12) on sensation and the cerebral cortex. Bartlett defined the schema as an active organization of past reactions, or of past experiences, which must always be supposed to be operating in any well-adapted organic response. Such schemata were thought to be ordered on a chronological basis and so constituted as to endow every fresh response, whether perceptual or motor, with appropriate skill or facility. In recognition, the scheme was conceived to operate *en masse* and to produce an orientation or attitude which made possible an appropriate differential reaction. In recall, on the other hand, the organism has in some sense "to turn round upon the schemata and to construct them afresh". This admittedly shadowy operation was attributed by Bartlett directly to the evolution of consciousness.

Many years ago, R. C. Oldfield and I (1942–3) made an earnest attempt to examine this theory and to trace the ways in which Bartlett's use of the concept of the schema differed from that of Head. I have no wish whatsoever to revisit this particular period piece. Nor do I think it would be helpful to review the valiant attempts made by Oldfield (1954), Broadbent (1958) and more recently Neisser (1967) to reinterpret some aspects of Bartlett's theory of schemata in the light of information theory and computer technology. These rescue operations are understandable but the theory, in my view never very plausible, is perhaps best forgotten.

How Constructive Is Recall?

Unlike his theory, Bartlett's experimental findings have stood up pretty well to the test of time. As Broadbent (1970*b*) has remarked, it is ". . . in some ways a measure of Bartlett's stature . . . that nobody seriously questions the factual results of his experiments". In a rough and ready way, they have been replicated over and over again in practical courses and laboratory exercises all over the world. Yet I feel there is a need to reassess Bartlett's treatment of recall as a constructive process and the reasons for his decisive rejection of the trace theory of memory.

My first doubts were sown many years ago, when, as one of Bartlett's research students, I was working on some problems of recognition and recall. I was able to show that if a figure which has previously been seen and reproduced a number of times is presented for recognition, along with other figures embodying changes which the subject has himself introduced in recall, only about half of these changes appear to have any influence on his subsequent recognition (Zangwill, 1937). I was later able to show much the same thing in the recognition of short stories (Zangwill, 1939). More recently, Prentice (1954) has produced striking evidence that "the process of recognition is unaffected by the very influences that make such dramatic changes in reproduction". Admittedly, he tested recognition very soon

after the subject had reproduced the material and with longer intervals, as in my experiments, greater interference might well have been anticipated. None the less, it is impossible to dissent from Neisser's conclusion that "experiments using the method of reproduction give only equivocal information about visual memory". (Neisser, 1967).

Another point to bear in mind is that both the material used by Bartlett and the circumstances in which recall was demanded in his experiments may well have conspired to produce greater inaccuracy than occurs normally. Gauld and Stephenson (1967) have recently repeated one of Bartlett's experiments on repeated reproduction using perhaps the best known of his stories, *The War of the Ghosts*. They have shown that strict instruction to avoid mistakes markedly reduces the number of errors made. They have also shown that subjects are in general fairly adept at picking out their own errors. It is probable, therefore, that subjects in experiments of this kind not uncommonly display an element of conscious invention and practise a kind of "gap-filling" by no means unusual in ordinary social intercourse. None the less, common experience suggests that the existence of unwitting transformation in recall, particularly over periods of time longer than those used in the Gauld and Stephenson experiment, can scarcely be contested.

In some experiments carried out several years ago by Dr B. R. Gomulicki, who worked with me at Oxford, we made a determined attempt to clarify this issue of the so-called constructiveness of recall (Gomulicki, 1953a, 1956). Our method was very simple: we selected 33 prose passages, all from modern English sources, of approximately similar difficulty but varying in length between 15 and 200 words. Each passage was read aloud to the subject, who reproduced it orally immediately afterwards. Analysis of the results showed that almost all the errors made under such circumstances were errors of omission, the process of recall being far better described as abstractive than as constructive. As we saw it, the elements of each passage were seriated in terms of relative importance, enough of the elements of least importance being eliminated to bring the content within the scope that can be encompassed in recall. This interpretation was strengthened by the finding that when we asked judges to try to distinguish between a précis written by one subject with an immediate reproduction of the same passage written by another, the resulting discriminations were little above chance level. That is to say, intentional and unintentional abstracts of the same material are virtually indistinguishable. I should also add that such experiments as we did on delayed reproduction suggested that the material, if it could be remembered at all, was given in increasingly abstracted form with the passage of time. Forgetting, therefore, is correlated with ever-increasing stringency of abstraction.

These findings led us to conclude that remembering is better described as an abstractive than as a constructive process, a point of view that was accepted soon after by Harry Kay (1955) on the basis of a very different type of experiment. This way of describing remembering does not of course mean that constructive importation never occurs—only too patently it does—or that social and cultural factors do not play an important role in governing both the matter and the manner of recall. Indeed abstraction, whether witting or unwitting, may itself be at least in part an acquired skill. But I think it would be proper to shift the emphasis from

reconstruction back to reproduction, even while accepting Neisser's *caveat* that precise repetition of any movement, any spoken sentence or any sequence of thought is extremely difficult to achieve. The basis of much long-term memory would seem to lie in an abstracted and highly edited verbal response to the general time-course of experience.

Conventionalization

I would like to dwell briefly on *conventionalization*, an idea which Bartlett took over from Social Anthropology and which as we have seen had an important influence on his choice of experimental method. Indeed his Fellowship Dissertation, which embodied the first results of his memory experiments, was subtitled "A Contribution towards an Experimental Study of the Process of Conventionalization". He also devoted a chapter in *Remembering* to the same topic.

As used by Bartlett, this rather inelegant term refers to the process whereby a cultural element, such as a work of art, transmitted by diffusion from one society to another gradually loses its representational significance and takes on a formal, conventional and non-representational character. The evil eye, for example, may be initially represented by a lifelike picture of an eye but end up as a mere dot surrounded by a circle: it has become a conventional sign. Bartlett hoped to compare the process of conventionalization as it is manifested in a group through influences initiated by contact with other groups with conventionalization as manifested in the individual through the operation of purely psychological factors. He hoped especially that his Method of Picture Writing would throw some light on this matter, though I am not convinced that it really did. At the same time, Bartlett (1916*b*) made some interesting introspective observations on the ways in which one's idea of a person or place may be embodied in a single, particular image, which comes to exhibit a certain dominance in mental life. It might be that if we understood better how conventional signs originate in the individual, we would be in a position to make a contribution of some importance to social psychology.

While there have been several studies designed to bring out the influence of social and cultural factors on memory, most of them using the Method of Serial Reproduction (e.g. Maxwell, 1936; Northway, 1936; Carmichael, 1939; Allport and Postman, 1948), I know of only one study in which a direct comparison was made between conventionalization in the individual and in society. This was an inquiry by T. H. G. Ward (1949), who worked with me for a time at Oxford, into changes in the design of some early coin types. Ward pointed out that typical coins struck in Macedon in the fourth century B.C. underwent striking simplification in design when struck in France in the second century B.C. and in various parts of Britain in the course of the first century B.C. He noted particularly that the laurel wreath on the head of Apollo on the obverse of the coin became detached and simplified while the wheel of his chariot on the reverse was preserved only as a decorative item. Serial reproduction experiments with Oxford undergraduates as subjects using these two details from the original coins showed changes very similar to those observed in the series of coin types. Although, as Bartlett himself said at the time, Ward was perhaps a little lucky, his work does at least suggest that

more careful study of the factors governing change in serial reproduction could throw an interesting light on the evolution of certain artistic and decorative forms.

Memory and Habit

Influenced perhaps by his early reading of Bergson (1911), Bartlett always drew a rather sharp distinction between memory, by which he meant the recollection of unique events, and habit, by which he meant the running off of an acquired chain of reactions.   Unlike the Behaviourist School, Bartlett had scant respect for habits, regarding them as essentially primitive, low-level and uneconomic, and indeed as modes of behaviour into which we tend to fall particularly when we are "tired, delirious or slightly intoxicated".   (Skills, on the other hand, particularly when athletic, Bartlett always held in high esteem!)   Now I have always felt that Bartlett treated habit rather scurvily and that his relegation of it to so low a level in the evolutionary hierarchy was unjustified.

Since William James, it has been widely held that the formation of habits is a signally effective manner of dealing with those aspects of the environment that remain relatively constant over time.   In amnesic states due to injury or disease of the brain, the capacity to form habits is disrupted almost as completely as is memory in its more personal and social aspects.   This suggests a common neural basis for these two aspects of memory function.   On purely psychological grounds, too, the supposed inferiority of habit has often been questioned.   In an early study, Oldfield and I (1938) showed that reproduction of verbal material in the early stages of learning by heart shows all the characteristics ascribed by Bartlett to constructive recall.   These, however, gradually fall away as learning progresses. In general, we saw no need to postulate a mechanism for the acquisition of rote habits fundamentally different from that operating in substance recall.   I have also been impressed by the fact that aphasic patients typically have much greater difficulty in learning material by heart than in recalling it "in their own words", i.e. abstractive recall is easier than rote recall.   This might suggest that rote learning represents, if anything, a relatively high-grade and specialized use of language. As Neisser (1967), has pointed out, what seems to be simple associative revival of earlier responses may actually prove to be a complex process of search and construction.

Nonetheless, the stereotyping effect noted in experiments on repeated reproduction does suggest that material, once abstracted, tends thereafter to be reproduced in substantially the same form, especially if the time intervals between successive reproductions are relatively short.   That is to say, what is reproduced on all occasions after the first is not the original but one's own reproduction of it.   In other words, a habit has been acquired.   This was neatly brought out by Kay (1955) in an experiment on repeated reproduction of short stories in which the subject was on each successive occasion allowed to re-read the original passage immediately after he had reproduced it.   But even this re-exposure of the original produced strikingly little effect on the accuracy of subsequent reproduction; there was little, if any, further learning or tendency to depart from the form of the initial reproduction.   In Mary L. Northway's homely phrase, what is reproduced is not the original but what the subject makes of it—his response if you like—and it was

this that she supposed Bartlett really had in mind in putting forward his concept of the schema (Northway, 1940). Further, as Kay has shown, this schema is virtually determined in its content by the initial reproduction.

Two Memories or One?

In recent years, the view that memory comprises two distinct systems, one concerned with short-term and the other with long-term storage, has gained a wide measure of acceptance. This view is supported by a very considerable body of evidence, some of which has been recently examined by Donald Broadbent in an admirable review article (Broadbent, 1970a). Although the dichotomy as such remains controversial, it cannot be denied that, in the last 20 years or so, much has been learned about the operation of memory in its short-term aspects. Paradoxically, this advance appears to owe more to the example of Ebbinghaus than to that of Bartlett, in spite of the fact that several of the foremost workers in this field are to be numbered among Bartlett's former pupils, among them John Brown, Christopher Poulton and R. Conrad, as well, of course, as Broadbent himself. Indeed it may be said with some pride that Cambridge has taken the lead in developing theoretical interpretations of short-term memory—or more accurately its converse, forgetting—and this work has deservedly won recognition.

I do not think Bartlett himself explicitly accepted the distinction between short- and long-term memory, at all events at the time at which he was himself active in research. Indeed one of the points that always used to worry me about his treatment of recall as constructive is that it evidently did not cover short-term memory, the *verbatim* reproduction of words or phrases falling within the so-called immediate memory-span. Indeed Gomulicki's experiments were originally planned to study whether the transition from short-term to longer-term memory with increase in passage length was abrupt, implying a dual mechanism, or gradual, implying a unitary mechanism. In fact, it turned out to be gradual, abstraction becoming increasingly pronounced with increase in passage length and hence in the time taken to register it. Even so, a degree of incipient abstraction was often noted in passages whose length fell well within the memory span, or only slightly exceeded it. Further, if the material presented odd or unusual features, it was promptly rationalized precisely in the manner described by Bartlett. I have given examples of this in a note on immediate memory published some years ago (Zangwill, 1956), pointing out that constructive change is by no means confined to long-term memory and might on occasion be seen in the reproduction of very brief passages after very brief intervals of time. I also pointed out that such instant distortion of memory, especially when achieved by selective omission, might seem to have much in common with the Freudian concept of repression.

In this note, I further argued that it is of some importance to distinguish between rationalizations prompted by issues of factual probability on the one hand and those governed by purely linguistic considerations on the other. While conventionalizations of fact are the more dramatic, the possibilities of distortion due to preferred linguistic usage might certainly seem to merit further exploration. Although this particular topic does not appear to have been pursued further, there has of course

been a great deal of work during the past 10 years on the relations between semantic and syntactical factors in verbal memory, much of it carried out by George Miller and his colleagues and owing its ultimate inspiration to Chomsky. For example, Mehler (1963) has shown in a neat experiment not only that the syntactical structure of a sentence markedly affects the ease with which it is learned, but also that errors in the immediate recall of sentences are overwhelmingly syntactical in nature. He has therefore been led to suggest that one typically encodes transformed sentences as kernel sentences plus some "mental tag" indicating the appropriate transformation to be made on recall. Whatever one may think of Chomsky's linguistics, analysis along such lines as these may well take us further in elucidating the complex relations between sense and syntax in verbal memory.

I do not know whether or in what way these somewhat off-beat observations of mine fit into the conventional framework of short-term memory studies. But in the study of short-term no less than that of long-term memory, it seems to me vital to bear in mind Bartlett's warning that the use of artificially simplified material is biologically unsound and leads to the neglect of important principles of organization over time. By and large, what one may call the verbal learning industry has paid scant attention to this stricture and operates for the most part with materials and methods which we students of Bartlett's in the thirties would have regarded as quaintly old-fashioned. Miller (1956), it is true, has recognized the importance of what he calls "chunking", but his treatment strikes me as descriptive rather than explanatory. Only Neisser (1967), who expressly acknowledges his indebtedness to Bartlett, has pleaded for a more realistic approach to the problem of organization in short-term memory.

Images and Words

The next problem I want to consider is the relation between images and words in long-term memory. In his book and elsewhere, Bartlett had a good deal to say about images and their functions and came to regard them as having been evolved largely to short-circuit past experience operating *en masse* or in the form of sequentially organized reaction patterns. "An image", he wrote, "is a device for picking bits out of schemes, for increasing the chances of variability in the reconstruction of past stimuli and situations, for surmounting the chronology of presentations" (Bartlett, 1927). Words, on the other hand, he thought of primarily as involved in *formulating* the qualities and relations implicit in perceptual experience, and thereby making possible its orderly communication. Ordinary remembering involves a combination of imagery and formulation, depending on the material, the circumstances, and, above all, the individual reaction type.

In *Remembering*, Bartlett discusses the functions of images at greatest length in connection with his Method of Description, which, incidentally, is the only one of his methods to make much call on spatial memory, in this case the direction of regard of the faces portrayed. He points out that visualizers tend to be more confident in their judgements than vocalizers, though by no means necessarily more accurate. Indeed they are particularly prone to importation and invention. Vocalizers, on the other hand, make in general far greater call on verbal analysis

and classification and are better than visualizers at retaining an order of presentation. These descriptions they commonly revive whenever recall is demanded. While Bartlett makes clear that few subjects make exclusive use of either medium of recall, it is plain that there is a very real difference between them.

I once tried to devise an experiment to throw light on this difference, though it was never published. I tried to think of a task which, so far as I could see, could be solved only through the use of visual imagery. This involved the identification of previously exposed visual patterns using a difficult multiple choice test. As a moderately good visualizer, I did not myself see how this task could possibly be carried out without the use of visual imagery. As it turned out, however, the most successful subject was a colleague, now a distinguished Professor of Psychology, who claimed that, except when dreaming, he had never experienced visual imagery. His success on my task appeared to be based wholly on rapid verbal analysis and application of the resulting formula to each of the items presented for recognition. Although at the time I found this result vaguely disappointing, at least it served to confirm Galton's old observation that there is no correlation between the capacity to evoke visual images and the accuracy of visual recognition. Recognition appears to be a process *sui generis*.

Today, we are beginning to understand more clearly the relation between images and words in the storage and retrieval of visual information. While it is well known that language bears a special relation to the activities of the left cerebral hemisphere, it is less well known that there is a comparable, if perhaps somewhat less outspoken, relation between higher visual processes and the activities of the right cerebral hemisphere. This evidence comes in the first instance from comparative studies of patients with unilateral brain lesions in which a number of neuropsychologists have been engaged for some years (cf. Newcombe, 1969). More recently, it has received striking confirmation from the findings in patients who have undergone operative section of the cerebral commissures for the relief of intractable epilepsy (commissurotomy). A number of these patients have been studied in very great detail by Dr Roger Sperry and his associates at the California Institute of Technology, where by kind invitation of Dr Sperry I was privileged to spend a month last summer. At the time of my visit, Dr Colwyn Trevarthen, now at the University of Edinburgh, was engaged in association with several colleagues in investigating the differential capacities of the two separated hemispheres in the perception and storage of visual input. From this work (Levy, Trevarthen and Sperry, 1972), it appears plain that, in the processing, storage and retrieval of visual information under conditions in which the two hemispheres are in competition, the right hemisphere regularly takes precedence over the left. This applies, of course, only in cases in which the material is non-verbal and is particularly evident if it is of a kind that cannot readily be specified in verbal terms, e.g. in the recognition of pictures of human faces. Furthermore, there is evidence that each hemisphere handles visual input in its own characteristic manner. In general, the right hemisphere proceeds by global impression and direct matching; the left, by sequential analysis of key features. This difference parallels almost exactly that between the visualizer and the vocalizer as described by Bartlett. In the normal individual, of course, we must assume that every act of perception is governed by

activities in both halves of the brain. None the less, we may perhaps see in direct visual recognition, and perhaps also in the evocation of visual images, activity governed primarily by the right hemisphere in precisely the same way as linguistic analysis is governed by the left. Here, perhaps, is the beginning of true neurology of perception and memory.

Memory and the Brain

This leads me to my last question—the relation between memory and the functions of the brain. In spite of his close intellectual link with Sir Henry Head, Bartlett never showed any great interest in the physiological basis of his schemata or, more generally, in the ways in which neurological inquiry might throw light on the problems of memory. Here I soon came to differ from him. It seemed to me that inquiry into the disorders of memory might well clarify many of the issues which Bartlett had raised in *Remembering* and perhaps add a neurological dimension to his theory of schemata.

I therefore turned to the study of patients with organic disorders of memory, first in Cambridge and London and later, during the war, in Edinburgh, where it was my good fortune to work in Professor Norman Dott's Brain Injuries Unit. Characteristically, Bartlett gave me his full support in this work in spite of the fact that, for him, clinical studies held little appeal. At first, I worked mainly with cases of alcoholic Korsakoff syndrome, attempting to specify the aspects of memory principally affected (Zangwill, 1941, 1950), and later, with cases of head injury, endeavouring to define the scope of traumatic amnesia and to chart the course of its recovery (Zangwill, 1943, 1964; Paterson and Zangwill, 1944). I would like to comment briefly on one or two things that I learned from this work and which do, I think, have some bearing on Bartlett's ideas.

As a pupil of James Ward, Bartlett had little use for Faculty psychology. Memory, he held, is just a convenient name for a number of related functions which can never be strictly separated one from the other. "None", he wrote, "can set a ring round Memory and explain it from within itself." Remembering, moreover, ". . . is not a completely independent function, entirely distinct from perceiving, imaging or even from constructive thinking". We know, too, that Bartlett regarded remembering as heavily dependent on attitudes and interests and hence, presumably, as closely linked with the organization of personality. How relevant are these views to the findings of pathology?

While defect of memory is, of course, frequently accompanied by intellectual loss and personality change, I was soon to learn that it may on occasion be present with surprisingly little impairment in other spheres of psychological function. Following encephalitis, for example, it is not uncommon to find that the patient shows virtually complete preservation of his perceptual, intellectual and linguistic skills, yet can retain virtually nothing of his current experience, at all events for more than a minute or so (Rose and Symonds, 1960). This grave defect has usually been described in terms either of loss of the capacity to establish long-term memory or as a failure in the transfer of information from a short-term to a long-term memory store. Other kinds of interpretation are, however, possible. But

the highly specific nature of the deficit indicates that memory, even if we no longer wish to consider it as a faculty, depends on some fundamental property of the brain —perhaps that christened by James *plasticity*—which can be greatly reduced or even abolished altogether in certain pathological states without appreciable involvement of other aspects of higher brain function.

Another problem relevant to Bartlett's theory is *retrograde amnesia*, i.e. loss of memory for events which occurred prior, say, to a head injury and at a time at which brain function was wholly normal. In view of the fact that retrograde amnesia is often reversible, in large part at least, it has been customary to ascribe it to a failure of retrieval rather than to actual destruction of memory traces (Russell, 1971). Now if, as Bartlett supposed, the organization of memory is largely dependent upon interest and emotion, one would imagine that such factors would likewise govern the restitution of memory. This, however, is not the case. The order in which memories return in recovery from head injury appears to be governed not by their psychological importance but by their relative recency, by their time order, suggesting that chronological sequence is in some way built into the long-term memory store (Whitty and Zangwill, 1966). While it is true that Bartlett accepted chronological sequence as important in the organization of schemata, he gave it little emphasis in relation to retrieval. Yet the facts of amnesia strongly suggest that no theory of memory can afford to disregard the temporal order of events or to dispense with the concept of consolidation as a function of time.

Finally, one may consider briefly the issue of cerebral localization. The earlier work of Lashley (1929) on brain mechanisms and intelligence in the rat and monkey led many psychologists of my generation to question the doctrine of cerebral localization of psychological function and to view loss of memory as governed by the extent of a brain lesion rather than by its location. Many clinicians, too, held a basically similar view and regarded defect of memory as of no value whatsoever in the localization of cerebral disease. We now know that this is far from true. Circumscribed amnesic syndromes occur only with lesions in certain relatively well-defined regions of the brain, more especially the medial parts of temporal lobes and certain thalamic and hypothalamic nuclei (see e.g., Whitty and Zangwill, 1966; Victor, Adams and Collins, 1971). It is perhaps no coincidence that among those whose work has helped to establish this important clinico-pathological correlation is another quondam student of Bartlett's, Dr Brenda Milner, whose work at Montreal with Dr Wilder Penfield has won high distinction.

So far we have been concerned only with localization in relation to long-term storage. But we now appreciate that some aspects of short-term memory, too, are affected by cerebral lesions of determinate location. We know, for example, that immediate verbal memory is commonly at fault in aphasia (Zangwill, 1946), and may also suffer in patients who have sustained operations on the left temporal lobe, even if there is no clinically apparent disorder of language (Milner, 1958). As a rule, there is also defect in verbal learning but a recent study has suggested that loss of immediate verbal memory may on occasion be present without failure of learning or defect of long-term retention (Warrington and Shallice, 1969). In my view, there is little doubt that short-term storage of verbal information is

closely linked with the activities of speech perception, which we know to depend heavily upon the integrity of the temporal lobes, in particular the left. It is therefore by no means impossible that psychological models of the kind envisaged by Broadbent (1958) and others will in due course find their embodiment in the anatomy of the brain. While the localization of a deficit is, of course, something very different from the localization of a function, at least we now know where to look.

Conclusion

I have now come to the end of my revisit to *Remembering*—the first, I may say, since some time in the late 30s. This exercise I thoroughly enjoyed, even if it is one that few of my contemporaries, apart perhaps from those who aspire to become historians of psychology, will probably wish to emulate. At the same time, it gives me great satisfaction that many students today are paying their first visit to *Remembering*, exactly 39 years after it appeared and now happily available in paperback. Many of them seem to find the experience far from unprofitable; indeed there are welcome signs of a renewed interest in Bartlett's work. In spite of its theoretical shortcomings and the doubts that have been cast on the general validity of some of the experimental findings, Bartlett's book has remained something of a classic in modern psychology. In its freshness of approach, subtlety of observation, many-sidedness of interest and felicity of style, Bartlett's *Remembering* has a special place in the annals of our subject and in the hearts of those of us who were his pupils. It is to me a great honour to pay tribute to the memory of its author.

References

ALLPORT, G. W. and POSTMAN, L. (1948). *The Psychology of Rumor.* New York: Henry Holt & Co.

BARTLETT, F. C. (1916a). An experimental study of some problems of perceiving and imaging. *British Journal of Psychology*, 8, 222–66.

BARTLETT, F. C. (1916b). Transformations arising from repeated representation: a contribution towards an experimental study of the process of conventionalization. Unpublished Fellowship Dissertation, St John's College, Cambridge.

BARTLETT, F. C. (1927). The relevance of visual imagery to the process of thinking. III. *British Journal of Psychology*, 18, 23–9.

BARTLETT, F. C. (1932). *Remembering: A Study in Experimental and Social Psychology.* Cambridge University Press.

BERGSON, H. (1911). *Matter and Memory.* London: Swan Sonnenschein.

BETZ, A. (1910). Vorstellung und Einstellung: I. Wiedererkennen. *Archiv für die gesamte Psychologie*, 17, 266–96.

BROADBENT, D. E. (1958). *Perception and Communication.* Oxford: Pergamon Press.

BROADBENT, D. E. (1970a). Psychological aspects of short-term and long-term memory. *Proceedings of the Royal Society, London, B*, 175, 333–50.

BROADBENT, D. E. (1970b). Frederic Charles Bartlett, 1886–1969. *Biographical Memoirs of Fellows of the Royal Society*, 16, 1–13.

CARMICHAEL, D. M. (1939). Psychology among the Eskimos. *The Eagle*, (St John's College, Cambridge), **51**, 115–31.

GAULD, A. and STEPHENSON, G. M. (1967). Some experiments relating to Bartlett's theory of remembering. *British Journal of Psychology*, **58**, 39–49.

GOMULICKI, B. R. (1953*a*). Recall as an abstractive process. Unpublished D.Phil. thesis. University of Oxford.

GOMULICKI, B. R. (1953*b*). The development and present status of the trace theory of memory. *British Journal of Psychology Monograph Supplement*, **39**, pp. 94.

GOMULICKI, B. R. (1956). Recall as an abstractive process. *Acta Psychologica*, **12**, 77–94.

HEAD, H. (1926). *Aphasia and Kindred Disorders of Speech*. Cambridge University Press.

HEAD, H. and HOLMES, G. (1911–12). Sensory disturbances from cerebral lesions. *Brain*, **34**, 102–254.

KAY, H. (1955). Learning and retaining verbal material. *British Journal of Psychology*, **44**, 81–100.

LASHLEY, K. S. (1929). *Brain Mechanisms and Intelligence: A Quantitative Study of Injuries to the Brain*. Chicago: University of Chicago Press.

LEVY, J., TREVARTHEN, C. and SPERRY, R. W. (1972). Perception of bilateral chimeric figures following hemispheric deconnection. *Brain*, **95**. In press.

MAXWELL, R. S. (1936). Remembering in different social groups. *British Journal of Psychology*, **26**, 30–40.

MEHLER, J. (1963). Some effects of grammatical transformations on the recall of English sentences. *Journal of Verbal Learning and Verbal Behavior*, **2**, 346–51.

MILLER, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, **63**, 81–97.

MILNER, B. (1958). Psychological defects produced by temporal lobe excision. *Proceedings of the Association for Nervous and Mental Disorders*, **36**, 244–57.

MÜLLER-FREIENFELS, R. (1913). Typenvorstellungen und Begriffe (Untersuchungen zur Psychologie des Denkens). *Zeitschrift für Psychologie*, **64**, 386–443.

NEISSER, U. (1967). *Cognitive Psychology*. New York: Appleton-Century-Crofts.

NEWCOMBE, F. (1969). *Missile Wounds of the Brain: A Study of Psychological Deficits*. London, Oxford: Oxford University Press.

NORTHWAY, M. L. (1936). The influence of age and social group on children's remembering. *British Journal of Psychology*, **26**, 11–29.

NORTHWAY, M. L. (1940). The concept of the schema. *British Journal of Psychology*, **30**, 316–25.

OLDFIELD, R. C. (1954). Memory mechanisms and the theory of schemata. *British Journal of Psychology*, **45**, 14–23.

OLDFIELD, R. C. and ZANGWILL, O. L. (1938). The acquisition of verbal repetition habits. *British Journal of Psychology*, **29**, 12-26.

OLDFIELD, R. C. and ZANGWILL, O. L. (1942–3). Head's concept of the schema and its application in contemporary British psychology. *British Journal of Psychology*, **32**, 267–86; **33**, 58–64, 113–129, 143–149.

PATERSON, A. and ZANGWILL, O. L. (1944). Recovery of spatial orientation in the post-traumatic confusional state. *Brain*, **67**, 54–68.

PHILIPPE, J. (1897). Sur les transformations de nos images mentales. *Revue Philosophique*, **43**, 481–93.

PRENTICE, W. C. H. (1954). Visual recognition of verbally labelled figures. *American Journal of Psychology*, **67**, 315–20.

RIVERS, W. H. R. (1912). Conventionalism in primitive art. *Reports of British Association for the Advancement of Science* (Section H), 599.

ROSE, F. C. and SYMONDS, C. P. (1960). Persistent memory defect following encephalitis. *Brain*, **83**, 195–212.

RUSSELL, W. RITCHIE (1971). *The Traumatic Amnesias*. London, Oxford: Oxford University Press.

VICTOR, M., ADAMS, R. D. and COLLINS, G. H. (1971).   *The Wernicke-Korsakoff Syndrome.*
    Oxford: Blackwell.
WARD, T. H. G. (1949).   An experiment on serial reproduction with special reference to
    the changes in the design of early coin types.   *British Journal of Psychology*, **39**, 142–7.
WARRINGTON, E. K. and SHALLICE, T. (1969).   The selective impairment of auditory-verbal
    short-term memory.   *Brain*, **92**, 885–96.
WASHBURN, M. R. (1916).   *Movement and Mental Imagery: Outlines of a Motor Theory of
    Consciousness.*   Boston, New York: Houghton-Mifflin.
WHITTY, C. W. M. and ZANGWILL, O. L. (Eds) (1966).   *Amnesia.*   London: Butterworths.
ZANGWILL, O. L. (1937).   An investigation of the relationship between the processes of
    reproducing and recognizing simple figures, with special reference to Koffka's trace
    theory.   *British Journal of Psychology*, **26**, 250–76.
ZANGWILL, O. L. (1939).   Some relations between reproducing and recognizing prose
    material.   *British Journal of Psychology*, **29**, 370–82.
ZANGWILL, O. L. (1941).   On a peculiarity of recognition in three cases of Korsakoff's
    psychosis.   *British Journal of Psychology*, **31**, 230–48.
ZANGWILL, O. L. (1943).   Clinical tests of memory impairment.   *Proceedings of the Royal
    Society of Medicine*, **36**, 576–80.
ZANGWILL, O. L. (1946).   Some qualitative observations on verbal memory in cases of
    cerebral lesion.   *British Journal of Psychology*, **37**, 8–19.
ZANGWILL, O. L. (1950).   Amnesia and the generic image.   *Quarterly Journal of Experi-
    mental Psychology*, **2**, 7–12.
ZANGWILL, O. L. (1956).   A note on immediate memory.   *Quarterly Journal of Experi-
    mental Psychology*, **8**, 140–3.
ZANGWILL, O. L. (1964).   Psychological studies of amnesic states.   *Third World Congress
    of Psychiatry.*   Montreal, 1961.   **3**, 219–22.

ERROR PATTERNS IN A CHORD KEYBOARD RESPONSE TASK

D. J. GUEST, M. E. SIME AND T. R. G. GREEN

MRC Social and Applied Psychology Unit, Department of Psychology, University of Sheffield

Errors in a serial choice-reaction time experiment employing two-finger chord responses to alphabetic stimuli are shown to be distributed in a significantly non-random fashion. Two factors are postulated which are relevant to prediction of the error response. These are the presence or absence of a common finger, and the difference in finger separation, between the actual and the required response. It is suggested that the chord response is coded in terms of a pattern (the distance between the two response components). Responses which mirror the correct response about the centre of the hand are also shown to be significantly more frequent.

Introduction

Only comparatively recently have the errors which occur in keyboard tasks been studied in detail. In a series of experiments Rabbitt (1968) has examined the latencies of errors and error-correcting responses and has shown that, in a highly compatible serial choice-reaction time experiment, error latencies tend to be faster than correct reaction times, and that errors can be corrected (by making the response which should have been made) much more quickly than correct responses can be made. He argues that the most economical explanation for errors which can be corrected in such tasks is that they result from mistakes in motor coding rather than from coding errors earlier on in the information processing chain.

Very little research has been carried out on the types of errors which occur in such tasks and their partitioning into distinguishable classes. Either errors in a reaction time task occur entirely at random, or their distribution will be non-random and may throw some light on the coding processes at work. A useful technique in the analysis of such errors is the confusion matrix, in which error responses are plotted against those responses which should have occurred. Hick (1952) gives a matrix of one subject's responses for a ten-choice task. Seibel (1962), in replicating an experiment by Ratz and Richie (1961) where subjects learnt all of the 31 possible chord responses on a five-key keyboard, showed that different chords gave rise to different error percentages, and that reaction times for particular chords were highly correlated with error probability. More recently Kornblum (1969) has shown significant differences in the various types of error which occur within both a discrete and a serial reaction-time situation. Rabbitt and Vyas (1970) describe an experiment similar to Kornblum's in which the differing incidence of error types is further reflected in differing response times.

The experiment to be described here was initially designed to investigate the effects of adaptively biased sequences of stimuli upon keyboard training, but the distribution of errors within the task is sufficiently interesting to be worth mentioning for its own sake.

Method

Apparatus

The stimuli were ten brightly lit and easily discriminable letters of the alphabet ¾ in. (19 mm) high, which appeared on a rear projection display at head height and at a distance of 2 ft (0·6 m). Responses were made on a five-key keyboard mounted at waist level. The keys were micro-switches and required a pressure of 200 g for closure. An Elliott 903 computer was used to present the stimuli and to record responses.

Subjects

Two groups of subjects were run, each consisting of five right-handed University students.

Procedure

Each subject received the same one-one correspondence between the 10 stimuli and the 10 possible two-key responses for a five-key keyboard. This correspondence remained invariant for the duration of the experiment. Subjects were not initially informed of the stimulus-response relationships (except that to the 10 stimuli corresponded 10 different two-key responses) and had to discover them by trial and error. Accordingly, a correction procedure was used; the stimulus would remain until the correct two-key combination was pressed when it was replaced by another stimulus within 10 msec. An audible click occurred on a loudspeaker as the stimulus changed. The subject could therefore know that he had made a correct response even when the same stimulus was re-presented.

Subjects were instructed to perform the task as fast as possible and not to worry too much about errors.

Recording

The keyboard was scanned at 10 msec intervals and a correct response was registered whenever the two correct keys and no other key was depressed. No constraint was placed on the time interval between the selection of the correct keys. The depression of any other key during this interval was recorded but did not invalidate the response. There was thus no upper limit on the indifference interval for making a chord. All changes of state of the keyboard were recorded. The RTs of correct responses were recorded accurate to the nearest 10 msec.

Practice

Subjects attended for six successive days and on each day performed four practice runs and one test run. Each run contained 200 stimuli, so that in all a total of 6,000 correct responses was collected for each subject. After each run was a rest period of about 2 min during which the subject was told the time he had taken for the run to the nearest second.

Stimulus sequence

A control group received a random sequence of stimuli throughout both practice and test runs. A second group, the Adaptive group, received a random stimulus sequence on the test runs, but was given an adaptively biased sequence of stimuli on the practice runs.

This bias was introduced as follows: Initially 10 stimuli were chosen at random (with replacement) from the 10 stimulus letters. These were then placed in a store from which the stimulus to be presented was selected at random. The time to make a correct response to this stimulus (RT) was measured and compared to a digital approximation of the median reaction time (the comparison time). If the RT was less than the current comparison time then the particular stimulus to which it referred was removed from the store and replaced by drawing randomly from the 10 stimulus letters. Otherwise the stimulus was allowed to remain in the store and in either case the RT was used to compute a new value of the comparison time. In this way stimuli to which the subject took longer than average to respond built up in the store and were thus presented more frequently. Subjects were not told that the task was Adaptive.

The letters used and response patterns corresponding to them were: C 00011; E 00101; F 00110; Y 01001; J 01010; L 01100; Q 10001; R 10010; T 10100; X 11000—and 1 = key depressed, 0 = key not depressed, and reading from left to right for thumb, forefinger etc. (The choice of letters was limited by hardware constraints.)

Results

In the analysis which follows, only two-finger responses were considered, and the first trial of each sequence of 200 was discarded.

Trials can be separated into instances where the response is immediately correct (IC), where a single error response is followed by a correct response (SE) or more than one error is followed by a correct response. On comparing performance on the test runs for the two groups, using Analysis of Variance with a repeated measures design (Winer, 1962), there was no significant difference between groups on the time scores or the number of IC trials per run. The time per test run declined significantly with practice for both groups.

In the test runs the control group performed the more slowly and also made fewer immediately correct responses. These differences were not significant and were attributable to the extremely poor performance of one subject in the control group. His error response distributions are included in the analysis since they did not seem to be atypical. For the error analysis which follows both practice and test runs were pooled.

For each group a 10 × 10 error matrix was compiled (see Tables I and II). For this purpose SE and IC trials only were used—this was to increase the proportion of trials during which subjects had at least partial knowledge of the code. In the matrices the rows represent the responses which should have been made and the columns correspond to the response actually made. The main diagonal from top left to bottom right therefore refers to immediately correct responses, while the entries in other cells represent the incidence of the particular confusions.

The total number of trials for each group was, after subtracting the first trial of each run, 29,850. The difference between this number and the grand totals on Tables I and II therefore gives the number of trials resulting in errors which were not corrected immediately, i.e. 2570 and 1589 for the control and adaptive groups respectively. These trials are not included in the error analysis which follows.

It is evident from a comparison of the range of the row totals on Tables I and II that the adaptive procedure did introduce bias in the frequency distribution of the stimuli. Furthermore, if we accept that the distribution of IC responses produced by the control group reflects the difficulty in learning these responses we would

TABLE I

Confusion Matrices (IC and SE response only)—Control group

Stimulus	Response required	Response made										Total
		00011	00101	00110	01001	01010	01100	10001	10010	10100	11000	
C	00011	2430	100	73	10	24	17	16	34	15	4	2723
E	00101	59	2288	141	25	37	53	4	12	21	5	2645
F	00110	97	97	2473	10	23	89	0	10	9	3	2811
Y	01001	7	12	6	2251	37	13	78	118	24	127	2673
J	01010	10	15	59	22	2505	75	4	42	20	22	2774
L	01100	5	21	112	15	151	2232	2	11	67	39	2655
Q	10001	3	3	6	57	11	7	2822	27	10	21	2967
R	10010	20	9	10	62	29	13	46	2321	158	17	2685
T	10100	3	15	20	24	27	75	5	120	2269	45	2603
X	11000	4	4	4	29	8	28	31	13	57	2566	2744
	Total	2638	2564	2904	2505	2852	2602	3008	2708	2650	2849	27280

Table II

Confusion Matrices (IC and SE responses only)—Adaptive group

Stimulus	Response required	00011	00101	00110	01001	01010	01100	10001	10010	10100	11000	Total
C	00011	2591	54	65	3	13	16	3	6	5	5	2761
E	00101	29	3111	107	22	60	14	10	6	30	3	3392
F	00110	13	46	2758	5	33	28	3	13	12	3	2914
Y	01001	2	12	1	2852	25	7	16	115	26	29	3085
J	01010	6	26	28	23	2672	47	2	31	26	4	2865
L	01100	4	13	73	15	50	2841	3	8	19	13	3039
Q	10001	3	1	3	39	4	3	2084	24	3	5	2169
R	10010	5	10	10	58	15	10	18	2616	53	13	2808
T	10100	1	24	12	23	32	24	4	39	2652	19	2830
X	11000	5	2	7	19	11	42	8	13	20	2271	2398
	Total	2659	3299	3064	3059	2915	3032	2151	2871	2846	2365	28261

Response made

expect some relationships between this distribution and the distribution of the row totals of the adaptive group. Spearman's rank correlation is significant ($r_s = 0.68$, $P < 0.05$), indicating that the more difficult items were in fact presented more frequently in the adaptively generated sequence. The bias was not very great and, as reported above, did not result in a significant difference in learning.

The difference in stimulus sequence for the two groups demands, however, a separate error analysis for each. The effects which we report below are apparent in both groups and are therefore sufficiently robust not to be masked by these differences.

Error Analysis

On examining the confusion matrix there appear to be two main factors which are relevant to predicting the incidence of a particular confusion given a required response. These are: (1) confusions appear to be more common where the actual (error) response shares a common digit with the required response; (2) a measure which we will call Delta and which refers to the difference in the separation of the fingers when comparing actual and required response. The two fingers which constitute a particular response may be separated by a gap of 0, 1, 2 or 3 fingers. The required response for Q (10001), for instance, has a separation of 3 while F (00110) has a separation of 0 and so on. If the required response was R (10010) while the actual response was X (11000) then the Delta measure for such a confusion would be modulus $(2 - 0) = 2$. Similarly, the Delta of T (10100) and Y (01001) would be modulus $(1 - 2) = 1$.

Those confusions with a Delta of 0 may also be sub-classified into mirror and other (non-mirror) responses. A mirror response is defined as the reflection of the correct response about the middle finger of the hand, e.g. the mirror response of F (00110) is L (01100). It will be noted that mirror responses may or may not share a common digit with the correct response, and that both Q and J mirror into themselves.

TABLE III

Mean error scores (corrected for number of cells at risk) classified under Delta measure and presence/absence of common finger

	0 (Mirror)	0 (Non-mirror)	1	2	3
Control group Delta					
Common finger	11·85	11·85	10·92	4·09	3·55
No common finger	9·40	3·20	2·40	1·83	0·75
Adaptive group Delta					
Common finger	7·75	6·65	5·34	1·91	0·95
No common finger	9·15	4·35	1·90	1·00	0·60

We can place all the errors in the confusion matrices into one of the ten categories defined by the two factors (i) presence/absence of common digit and (ii) the four values of Delta, 0, 1, 2 and 3, not forgetting that a Delta of 0 is further split into mirror and non-mirror response. In Table III the mean numbers of errors in the 10 classes made over all six days for the two groups are presented. Error responses with a common finger are significantly more common than those without (Wilcoxon matched pairs, $P < 0.01$, two-tailed). Furthermore, confusions between two responses become less frequent as the Delta measure increases.

Table IV presents the average percentage of errors made in the four stimulus classes composed of differing "gap" values, i.e. separations between the two fingers of the response. Excluding the class with a gap of three (which contains only one member, Q) no significant differences exist between the other classes (non-parametric analysis of variance). This observation indicates that the effects of Delta cannot be due to a simple addition of the effects of each gap value, but must presumably be due to a tendency to reproduce in the error response the gap value of the correct response even when the placement is lost.

TABLE IV

Errors made in the four response classes as a percentage of the total number of presentations for each class

Control group				
Gap	0	1	2	3
Percentage	13·4	14·2	14·0	6·6

Adaptive group				
Gap	0	1	2	3
Percentage	6·1	8·4	8·5	3·2

A test of mirror errors against non-mirror errors (both with a Delta of 0) shows that, for confusions with no common finger, mirror errors are significantly more frequent in both groups (Wilcoxon matched pairs test, $P < 0.02$, two-tailed); but where a common finger is present, though mirrors are more frequent they are not significantly so.

The performance of the subjects on days 1–3 and days 4–6 taken separately showed that the Delta measure effects were reliable over practice.

Table V shows the mean frequency of Delta = 0 errors occurring with different

TABLE V

Mean number of errors per subject associated with a displacement of the response to left or right of the required position whilst retaining the pattern intact (i.e. Delta = 0). The data are corrected for the number of cells at risk

	Distance displaced		
Group	1	2	3
Adaptive	9·184	2·800	1·000
Control	11·950	2·166	0·800
Mean	10·567	2·483	0·900

amounts of displacement from the correct response position. The data are corrected for the number of ways each displacement can occur.

Discussion

This experiment was not designed with response error analysis as its primary objective. No provision was made, therefore, for removing systematic effects due to stimulus, as opposed to response, confusions. This could have been done by randomizing SR mappings over subjects. Despite the presence of noise introduced by stimulus confusions the results indicate clearly that confusion frequencies were strongly affected by three factors relating the required response to the response erroneously produced. Furthermore, the effects can be seen to be common to both groups examined and were not influenced in any obvious way by the different treatments. We therefore assume that they may be easily repeated.

The main factors relating error to required response were: a low Delta value, the existence of a common finger, and the presence of mirroring between the correct and actual response.

It will be recalled that mirroring was defined as reflection about the centre of the keyboard (or hand). Also that Delta refers to the difference in the separation of the fingers when comparing actual and required response. Proximity of placement of actual to required response can be added to the above factors.

The observed high frequency of errors having a component in common with the correct response would be consistent with a simple coding model where the two components were independently coded relative to some anchor point. In such a system the probability of both components being wrong would be the product of the probabilities for each. This statement of course assumes that correction has been made for the various constraints imposed by the response space. Such a model would predict that Delta zero errors would be less frequent than the sum of errors where Delta was non-zero, since Delta zero errors require that both response components are displaced. In fact our data shows clearly (Table III above) that Delta-zero errors are more common than the sum of the errors for other Delta values, being more numerous in the ratio 1·8 : 1. This, together with the decreasing trend in error frequency as Delta increases, suggests a strong interdependence in the placement of response components, i.e. the response is coded as a pattern in the keyboard space.

A simple model which might be entertained to take account of this could be one where the distance of the first component from an anchor point is stored, together with the distance of the second component from the first. The anchor point could well be a keyboard extremity, so that T (10100) could be stored as some analogue of the number pair 0, 2 or 2, 2 (depending on the anchor point chosen) and Y (01001) as 1, 3 or 0, 3. In this case corruption of the first member of the stored pair would result in a placement error but (response space permitting) retain the pattern. Corruption of the second member only would result in a common component error. Such a model would run into some difficulty with mirror errors as they could only be explained by a switch of anchor points.

As the centre of the hand and the centre of the response space are coincident

on a five-key keyboard our results cannot indicate whether the axis of rotation in mirroring a response is anatomically or spatially determined. Further experimentation is required to resolve this issue.

Our results indicate that the pattern of the required response persists with some tenacity in the error responses (Table III) and that small displacements from the correct position on the keyboard are much more likely than large displacements (Table V). Clearly the most probable error in this task is that of displacing the entire response pattern a single keyboard position to left or right.

It is interesting to consider whether there is an analogy here with the adjacent errors in single finger tasks. An adjacent error, being the selection of the key immediately to the left or right of the required key, is by far the most common error type in these tasks; cf. the error matrix presented by Hick (1952). Rabbitt (1968) has presented plausible arguments, on the basis of error correction times, that errors showing the properties of adjacent errors could be motor coding errors. (For a discussion of this see Green, Sime and Guest, 1972.) To sustain such an analogy we would need to show that the error correction times for displaced chords behave as do correction times for single finger responses. This we cannot do from our present data but the matter should prove worth further investigation. The displacement of a chord is less easily viewed as a peripheral slip of the motor system than is the displacement of a single finger response and might invite the view that motor encoding is a fairly high level process.

We believe it will prove informative to relate findings from chord response tasks to the much more fully explored single finger tasks. This will, however, call for more detailed time recording than was provided in this experiment. Useful information could have been gained, for example, had we recorded RTs for individual response components, allowing us to discover whether the classes of error which we found to differ significantly in incidence also differed in their response times.

With the increasing availability of on-line computers for this type of work, recording of complete RT information in complex response tasks has become practicable and was omitted in this instance only because the primary objective of the experiment did not require it. The authors now regret the omission.

References

GREEN, T. R. G., SIME, M. E. and GUEST, D. J. (1972) Some types of error in a choice-response task. *Quarterly Journal of Experimental Psychology*, **24**, 142–158.

HICK, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, **4**, 11–26.

KORNBLUM, S. (1969). Sequential determinants of information processing in serial and discrete choice reaction time. *Psychological Review*, **76**, 113–31.

RABBITT, P. M. A. (1968). Three kinds of error signalling responses in a serial choice task. *Quarterly Journal of Experimental Psychology*, **20**, 179–88.

RABBITT, P. M. A. and VYAS, S. M. (1970). An elementary preliminary taxonomy for some errors in laboratory choice RT tasks. *Acta Psychologica*, **33**, 56–76.

RATZ, H. C. and RITCHIE, D. K. (1961). Operator performances on a chord keyboard. *Journal of Applied Psychology*, **45**, 303–8.

SEIBEL, R. (1962). Performance on a five-finger chord keyboard. *Journal of Applied Psychology*, **46**, 165–9.

WINER, B. J. (1962). *Statistical Principles in Experimental Design.* New York: McGraw-Hill.

SOME TYPES OF ERROR IN A
CHOICE RESPONSE TASK

T. R. G. GREEN, M. E. SIME AND D. J. GUEST

MRC Unit for Social and Applied Psychology, Department of Psychology, University of Sheffield

Errors from a serial response task involving single-finger responses to alphabetic stimuli are analysed and discussed in relation to findings which have been reported from tasks with more compatible stimulus-response relationships. Errors are divided into three distinguishable subsets and in each case found to have longer latencies than correct responses. Those which result from mirroring the required response about the centre of the hand are found to resist elimination during practice and their frequency seems to depend on the type of code used. In all cases error correction times are faster than the times to make a correct response but mirror errors and errors involving a finger adjacent to the correct response are corrected faster than other errors. The findings are discussed in relation to the theory of choice reaction time and error correction.

Introduction

Present choice reaction time models differ considerably in their postulates about the sources of error. Fitts (1966) proposed a Bayesian random-walk model in which sequential samples are drawn from noisy data: with each sample odds favouring stimuli are revised until a criterion (critical odds-value) is reached and a response is made. Errors, in this model, are caused by stimulus misidentifications, and their frequency is controlled by the criterion value. Laming (1968) has proposed a version in which errors can be caused by premature sampling. Somewhat similar models have also been described by Stone (1960) and Edwards (1965). All these classes of model predict that error reaction times (RTs) will be distributed in the same way as correct RTs.

Hick (1952) proposed a serial-classification model in which the stimulus is classified by following a decision tree. This model, together with slight variations such as those of Welford (1960) and Bertelson (1963), supposes that subjects under time pressure make errors by abandoning the final stages of classification and choosing at random among those stimuli not yet eliminated; error RT will then tend to be faster than correct RT.

One might remark that both these models are more like models of pattern discrimination than models of the whole reaction time process. In both cases, all or virtually all the errors are ascribed to failures in stimulus identification; neither deals explicitly with practice effects.

First attempts were made to test the models by examining error RTs, which are usually found to be faster than correct responses (Rabbitt, 1966; Schouten and

Bekker, 1967; Egeth and Smith, 1967; Laming, 1968; Hale, 1969) although Fitts (1966) observed no difference. Faster errors were originally taken to support the serial-classification model rather than the random walk model (at least in Fitts' version) and were not taken to indicate that errors were different in kind from correct responses (Egeth and Smith, 1967). Schouten and Bekker (1967), however, not only devised a reliable technique for forcing subjects to respond fast and thereby increase their errors, but also showed that the ratio of response speed to probability of being correct was constant over all percentages of errors; this is not predicted by either model. Rabbitt and Vyas (1970) continued the line of thinking and extended the analysis to show that there appeared to be a critical speed for a well-practised subject such that responding faster greatly increased error probability while responding more slowly did not greatly improve accuracy. On the basis of observations at various levels of practice, they put forward a tracking model in which the subject first has to buy information about his critical speed, by making errors, and then has to track that speed as his performance improves with practice. In the same paper, they also showed that the tracking model was over-simple, and that to provide better models a great deal more knowledge was required.

The tracking model assumes that subjects can usually detect and correct errors without external feedback. Although there is some support for this, it also means that the errors are likely to be motor errors, for it was shown that correction times for detectable errors were less than mean correct RT (Rabbitt, 1968) and that this is more easily accounted for if we assume that the errors are not failures of perceptual classification but of motor execution. The tracking model therefore reverses the emphasis of the two stimulus classification models, by taking the majority of errors away from stimulus misclassification and putting them into response selection and execution. It also predicts, therefore, that arbitrarily high accuracy is not achievable merely by lowering the mean response speed; once past the critical speed, further speed reductions will not help and only practice will reduce errors further. Rabbitt and Vyas (1970) argue that what is now necessary is a detailed examination and taxonomy of error types.

Although the tracking model explicitly considers the role of practice, it does not consider the effects of the translation component in mapping from stimulus to response. A variety of codes have been used: Fitts (1966) used lights paired spatially to the corresponding keys, and Hick (1952) used lights arranged in a circle; Hale (1969) used the digits 1, 2 and 3 requiring "key presses by the first three fingers of the right hand in direct correspondence with the stimuli", and Rabbitt (1968) also used digits, although he mapped either two or four stimuli onto one response. These tasks all need relatively little translation from stimulus to response. Guest, Sime and Green (1972) used a two-finger chord response task with alphabetic stimuli, requiring therefore a high level of translation, and found three factors contributing to errors. Two responses were more likely to be confused if they had a finger in common; if the separations between the fingers were similar; and if one response was the reflection of the other about the middle finger. The last of these, "mirror" errors, can conceivably occur in single-finger response tasks, but it is not clear where they would be placed in the "elementary taxonomy" of Rabbitt and Vyas (1970).

The experiment to be reported was designed to investigate error types in a 10-choice task with a high translation component, using single-finger responses. The method of analysis was to cast all responses in a confusion matrix and to partition the matrix into subjects. On the assumption that motor confusion errors (Rabbitt and Vyas, 1970) would tend to involve confusion of adjacent fingers, one subset consisted of all errors immediately adjacent to the correct response. One subset consisted of the mirror errors, involving confusions between thumb and little finger and between ring and index fingers. The third subset consisted of all other error made with the correct hand. These subsets exhaust all possible correct-hand errors; wrong hand errors were kept low, and response repetition effects also kept low, by requiring hands to be used alternately.

The tracking model makes predictions about the role of practice which were tested to some extent by using low levels of practice, so that the rate of elimination of error types could be examined. Finally, the role of the stimulus-response mapping was investigated by employing three different mappings.

Method

Apparatus

An Elliott 903 digital computer was programmed to present a random sequence of stimuli. Random generation was accomplished by mapping the integers 0–4 onto the five stimuli possible on any trial, and using the subject's last response time (in hundredths of seconds) *modulo* 5 to define his next stimulus. No discernible bias was detected in the sequences of stimuli, using frequency and runs tests.

The stimuli were letters of the alphabet displayed on a rear-projection digital display tube giving very discriminable stimuli of constant height and constant apparent distance from the subject. Filament bulbs were used, slightly over-run to reduce rise-time; no special precautions were taken since comparisons were intended to be relative, rather than absolute. The subject sat alone in an experimental room with the fingers of both hands resting on the corresponding keys of the 10-key keyboard, similar in design to that used by Fitts (1966), mounted in front of him at a comfortable height. Keys were of the micro-switch type with a making pressure of approx. 200 g. Beside each hand was a small red cue lamp.

Procedure

The stimuli were mapped 1 : 1 onto the 10 possible single-key responses. A correction procedure was used, i.e. the subject had to carry on responding until he made the correct response. A new trial would then start and the next letter would be displayed. Stimulus change occurred within 10 msec of correct response being made. Each session comprised 10 runs of 50 trials; runs were separated by a rest period of about 1 min, during which the subject was told the total time taken for the previous run in sec.

Subjects were instructed to use single key presses only, and were told that multiple-key presses would not be taken as correct even if one key were the correct response. After each correct response it was necessary for all keys to be released for at least 10 msec before the program would accept another response as correct.

The translation code was not revealed to the subjects, except for the information that a 1 : 1 mapping was employed, and they were therefore forced to discover the code initially by trial and error. The instructions placed mild stress on speed rather than accuracy: "don't worry too much about errors, but go as fast as you can".

Stimuli

One group of 12 subjects was run using ten stimuli, each paired with a unique response. In this group (Group 10S), stimuli were chosen so that the hands were used alternately; subjects were informed of this fact, and the two small red cue lamps were used to remind them which hand was appropriate on each trial by lighting the lamp beside that hand. A constraint was placed upon the sequence such that on no trial was the correct finger the same as it had been, ignoring hand change, on the previous trial.

Two further groups were run using five stimuli, each paired with one response on each hand. Hands were used alternately, as in Group 10S, and the same cue lamps were used: thus on any trial only one correct response was possible. In Group 5S–AN, each stimulus was paired with identical fingers, e.g. with the two ring fingers; in Group 5S–SP the pairing was reversed so that one stimulus was paired with left little finger and right thumb, one with left ring finger and right index finger etc. (The mnemonics AN and SP refer to the anatomical and spatial methods of pairing two responses.) No stimulus repetitions were allowed to occur in either of these groups.

The sets of letters used were: CFLQYXEJTR for the 10-stimulus group, and CFLQY for the two five-stimulus groups. These stimuli were randomly assigned to the responses for each subject, using drawings from Fisher and Yates' (1963) tables of random permutations. This procedure randomizes out systematic error effects due to stimulus confusion.

Subjects

Subjects were paid volunteer University students, each of whom took part in one session. None had to our knowledge taken part in any similar experiment.

After each subject had completed his session, his results were screened before being accepted for analysis. For the screening procedure two information-theoretical measures were employed. The responses of each run were cast in a 10×10 confusion matrix, response required, S, versus response given, R (using only the first response of each trial), and the measure T_{SR}, transmitted information, and $H_{(S)R}$, noise added by the subject, were computed. The criterion for accepting a subject was that T_{SR} exceeded $H_{(S)R}$ on at least two of the 10 runs; if this criterion was not met, the subject was replaced. The screening resulted in replacing one subject in Group 5S–SP.

Results

Each subject's first run was discarded, together with the first trial of all subsequent runs. On each remaining trial, the subject could make an immediately correct response; a single error followed by a correct response; or more than one error followed by a correct response. Immediately correct responses accounted for 56·2% of all trials analysed, single errors for 17·9%, and multiple errors for 25·9%. For the purposes of this paper, attention will be restricted to immediately correct and single-error trials, since this will increase the proportion of trials analysed on which subjects at least partially knew the translation code. Furthermore, except for ascertaining overall error-rates of groups, only errors occurring on the correct hand will be analysed. (Very few wrong-hand single errors were made, about 0·3% of all trials.)

The relative frequencies of mirror errors (ME), adjacent errors (AE), and other correct hand errors (OE) were obtained by totalling for each class and dividing by the number of ways each error could be made (respectively 8, 16 and 16). Table I shows these relative frequencies per run, averaged over runs 2–5 inclusive and runs 6–10 inclusive.

TABLE I

Mean frequency per run of each error class and of immediately correct responses, during early and late practice (adjusted for number of ways each can occur)

Practice	Group	Correct	ME	AE	OE
Early	10S	30·03	2·10	4·78	0·78
(Runs 2–5)	5S–AN	34·84	4·90	4·09	0·89
	5S–SP	39·90	4·35	3·63	0·81
Late	10S	39·63	2·05	3·70	0·48
(Runs 6–10)	5S–AN	47·89	4·88	3·16	0·76
	5S–SP	53·53	4·25	2·35	0·40

MEs clearly account for a large proportion of all errors; equally clearly they do not reduce with practice at a rate even approaching that of AEs and OEs (see Fig. 1). Tests were made of the rates of decline of the error types between early
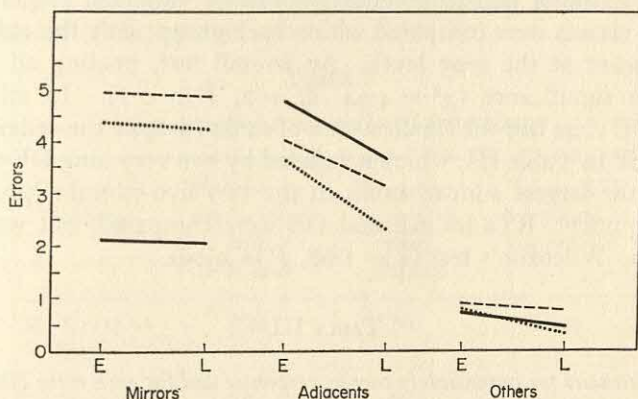

FIGURE 1. Data from Table I showing effect of practice on each type of error. E=runs 2–5; L = runs 6–10. —— 10S; ------- 5S–AN; ·········· 5S–SP.

and late practice, pooling all groups to obtain maximum sensitivity: these showed that the decline of MEs was not significant ($P = 0·488$, one-tailed, Wilcoxon) whereas AEs were significant at $P = 0·0008$ (one-tailed, sign test) and OEs were significant at $P = 0·013$ (one-tailed, Wilcoxon). The ratio of MEs to all other errors (including wrong-hand errors) is significantly greater in the two five-stimulus groups ($P < 0·015$, Mann-Whitney, two-tailed). Furthermore, these ratios do not bear any monotonic relationship to overall performance (Table II).

TABLE II

MEs as a percentage of all errors

Group	%ME	Total errors
10S	16·02	931·44
5S–AN	34·80	1012·40
5S–SP	34·30	901·56

Thus we must conclude that mirror errors are not only comparatively resistant to elimination, but are also more frequent at any given overall error-rate, in the five-stimulus groups.

It is not of course surprising that mirror errors should be more frequent in the 5S groups. In group 5S–AN a mirror error involves the use of the finger with which the stimulus is paired on the other hand. In group 5S–SP the spatial position, e.g. second key from the left, is the one which would be correct were the other hand being used. It is perhaps interesting that both the spatial and the anatomical pairing seem to produce errors in about equal measure. Perhaps only the 10S group can be held to show genuine "mirroring". In this case the stimuli for one hand have no reference on the other hand. It must be emphasized, however, that the resistance of these errors to elimination applies equally to all three groups.

Mean response times for each error class were computed, together with mean times for immediately correct responses (Table III).

Using the Friedman non-parametric analysis of variance, response times for the three error classes were compared within each group; only the 10S group gave a result significant at the 0·05 level. An overall test, pooling all groups, also failed to reach significance ($\chi^2 = 4·34$, $df = 2$, $P > 0·1$). In all three cases, however, the OE class had the smallest sum of ranks (despite the ordering of means for group 5S–SP in Table III, which is inflated by two very long OE times). The AE class had the largest sum of ranks in the two five-stimulus groups. For a subsequent argument, RTs for AE and OE were compared, but were again not significant using Wilcoxon's test ($z = 1·68$, $P = 0·09$).

TABLE III

Mean response latencies for immediately correct responses and for each error class, in 1/100 sec

Group	Correct	ME	AE	OE
10S	105·68	125·52*	119·82*	114·80
5S–AN	122·84	134·52*	147·10†	124·86
5S–SP	119·48	135·50*	147·90†	147·62†

* $P < 5\%$. † $P < 1\%$. All comparisons are against latencies for immediately correct responses. (Wilcoxon, two-tailed).

Separate comparisons were made, within each group, of the mean response time for each error class against the time for immediately correct responses. It can be seen from Table III that in all nine cases error times are longer, and that seven of the comparisons are significant at the 5% level or better (Wilcoxon, two-tailed), leading to the conclusion that in general errors take longer than correct responses. This is not due to an artifact arising from the sampling bias towards early practice, where errors are common and all response times are long, for essentially the same results were found when comparisons were made against only those immediately correct responses that directly preceded a single-error trial.

<div align="center">TABLE IV</div>

<div align="center">Mean times to correct errors of each class, in 1/100 sec</div>

Group	ME	AE	OE
10S	48·4	47·9	65·2
5S–AN	52·1	59·9	82·9
5S–SP	60·0	58·0	79·5

Finally, correction times were examined (Table IV). It is noticeable that the OEs take longest to correct in every group, while there is little to choose between MEs and AEs. Friedman's test showed that the correction times are significantly different from each other in Groups 10S and 5S–AN ($P < 0.01$ in both cases) and nearly reach the 5% level for Group 5S–SP ($\chi^2 = 5.64$, $df = 2$, $P = 0.06$). This indicates that the correction times for OEs are significantly long; reference to Table IV shows that they take about 200 msec longer to correct than MEs or AEs. A summary of the results is given in Table V.

<div align="center">TABLE V</div>

<div align="center">Summary of characteristics of error classes</div>

Error class	Frequency ranking	Frequency interaction with S–R code	Frequency reduction with practice	Error RT compared to correct RT	Correction time compared to correct RT
Mirror	Moderate to high	Yes	No	Slow	Fast
Adjacent	High	No	Yes	Slow	Fast
Other	Low	No	Yes	Slow: possibly faster than ME and AE	Fast, but slow compared to ME and AE

Discussion

There is one obvious point of divergence between our results and the result of low-translation tasks: in our case, errors take longer. At first blush that can be put down to long indecision latencies, a possibility to be considered later. Analysis of the confusion matrices has shown that three types of correct-hand response can readily be distinguished in terms of their frequencies, correction times, rate of elimination during practice and the interaction with the stimulus-response code; there is also a suggestion of a slight difference in error RTs.

Hick (1952) published the confusion matrix for one subject in a task with a lower translation component; the stimuli were 10 lights arranged in an approximate circle. We have re-analysed these results under the same headings as have been used

above. Little weight can be put on the results of one subject, of course. Furthermore, Hick did not disallow hand repetition, so that wrong-hand errors are more frequent than in our experiment; one would also prefer to be able to perform separate analyses according to whether the correct hand is repeated or alternated. Nevertheless, it is interesting to compare Hick's data (see Table VI) with the summary of ours given in Table V. In Hick's task there is no evidence that MEs occur more frequently than OEs, but AEs are very common; also error RT is longer for AEs than for OEs, even though all error RTs are shorter than correct RTs.

TABLE VI
Data extracted from a confusion matrix for one subject published (Hick 1952)

	N	RT	N per cell	Cells
Correct	1415	0·447	141·5	10
ME	9	0·3756	1·125	8
AE	408	0·4186	25·50	16
OE	22	0·3536	1·375	16
WH	76			
Total	1930			

We have partitioned the data in accordance with the error categories discussed in this paper.

In the taxonomy proposed by Rabbitt and Vyas (1970) the available candidates are motor confusion errors and perceptual misclassifications. The natural hypothesis is that the adjacent errors are motor confusion errors, and that the "other" errors—those correct-hand errors which are neither adjacent to, nor mirrors of, the correct response—are due to mistranslating the stimulus or to ignorance of the code. The correction times support that hypothesis: AEs are corrected about 200 msec faster than OEs. Rabbitt (1968, p. 187) argued that, in non-correction tasks, "the most simple hypothesis which, without any modification, covers all the results known to the author . . . is that subjects only detect their motor errors, and that they do this by comparing feedback from a response against a motor programme which remains available after the response has been made". He goes on to qualify the applicability of this comment, particularly with reference to practice effects, but his hypothesis as it stands is in accord with the present data on AEs and OEs.

The identification of AEs with motor confusions, and OEs with mistranslations, is also reasonably well supported by the other aspects of the data. Frequencies of both types decrease with practice, as would be expected, and their frequencies do not interact with the stimulus-response code used. In view of the high translation component of the task, it would also seem to follow that if OEs are mistranslations they should be associated with long indecision latencies. Response time for OEs would then be longer than the correct RT, which is a result observed, albeit weakly. On the other hand, if AEs are motor confusions, they should not be associated with long RTs; whereas they are in fact significantly longer than correct RTs.

If the indecision hypothesis is to be retained, it is necessary to account for long AE times. One way to do this is to postulate that response selection is a mechanism operating with the same speed-accuracy trade-off as has been postulated for stimulus identification. When the subject is unsure of the translation, his indecision latency will rise and the probability of mistranslation will also rise. Even if he translates correctly, he will attempt to compensate for the long indecision latency by hurrying the response selection, thereby increasing the probability of a motor confusion. But, although that hypothesis accounts for the long error times found in this experiment, it fails to account for the presence of AEs in Hick's low translation task, in which all error RTs were shorter than correct RTs. It would be necessary to postulate that in the response selection the subject grossly overcompensates for his excessive translation time. A further difficulty is the slender evidence that response times for AEs are actually longer than OEs, which if supported would be quite inexplicable under the trade-off hypothesis. In the event it seems difficult to preserve the "obvious" explanation of the long error RTs in our experiment while making any simple identification of any error class with motor confusions or with mistranslations.

The mirror errors are still more paradoxical. They cannot plausibly be put down to mistranslations, and the fast correction times associated suggest that they are motor errors. On the other hand their frequency interacts with the type of stimulus-response code used, being much higher in the two five-stimulus groups than in the 10-stimulus groups; they were found in a chord response task with alphabetic stimuli (Guest, Sime and Green, 1972); they were not found by Hick (1952). All this suggests that mirror errors are a more central phenomenon than a straightforward motor error. One possibility is that some sort of basic mirror factor exists; the fingers could, for example, be coded in terms of their distance from the middle finger, together with a sense indicator coded separately. Losing the distance code would then result in an OE or AE while losing the sense code would result in an ME. The probability of losing the sense code could quite well interact with the degree of work involved in translating from stimulus to response, which would explain the relation of mirror error frequency to stimulus-response mapping. Since OE frequency does not show the same relation it would be necessary to postulate that the probability of losing the distance code did not interact with translation difficulty: why it should be unlike the sense code in this respect is difficult to account for, a failing that makes the whole hypothesis suspect. Furthermore, the predictions it makes in the case of chord responses do not fit with the result of Guest, Sime and Green (1972). If the two components of a chord response are coded separately, each as a distance code and a sense code, it is to be expected that the errors due to one sense code will be more frequent than mirrors of the whole response, caused by reversing both sense codes; just the opposite was found.

The results of close examination of errors in a multiple-choice task have therefore been to discriminate three classes of errors; these appear from their characteristics to have different aetiologies, none of which is readily accounted for by existing theories of the choice reaction process. Nor is this because our experiment, by using relatively unpractised subjects in a high translation task, falls outside the domain of such theories, as was argued in the introduction. Our present assess-

ment of existing theories is that those which ascribe errors to perceptual mis-classification fail to account for the largest proportion of errors, the AEs, and that the error taxonomy of Rabbitt and Vyas (1970) needs considerable enlargement: but at present we cannot offer convincing hypotheses about the nature of the error types we have described.   It seems most unlikely that progress in this area can be made without many further attempts to classify and explain different error types in a wide variety of tasks and practice levels, a conclusion echoing Rabbitt (1968). Perhaps we should not be surprised that the simplest hypotheses seem to be in-sufficient.   Sophisticated mechanisms make sophisticated mistakes.

References

BERTELSON, P. (1963).   S–R relationships and reaction time to new versus repeated signals in a serial task. *Journal of Experimental Psychology*, **65**, 478–84.

EDWARDS, W. (1965).   Optimal strategies for seeking information: models for statistics, choice reaction time, and human information processing. *Journal of Mathematical Psychology*, **2**, 312–29.

EGETH, H. and SMITH, E. E. (1967).   On the nature of errors in a choice reaction task. *Psychonomic Science*, **8**, 345–6.

FISHER, R. A. and Yates, F. (1963).   *Statistical Tables for Biological, Agricultural and Medical Research*.   Edinburgh: Oliver and Boyd.

FITTS, P. M. (1966).   Cognitive aspects of information processing: III.   Set for speed vs. accuracy. *Journal of Experimental Psychology*, **71**, 849–57.

GUEST, D. J., SIME, M. E. and GREEN, T. R. G. (1972).   Error patterns in a chord keyboard response task. *Quarterly Journal of Experimental Psychology*, **24**, 139–48.

HALE, D. J. (1969).   Speed-error tradeoff in a three-choice serial reaction task. *Journal of Experimental Psychology*, **81**, 428–35.

HICK, W. E. (1952).   On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, **4**, 11–26.

LAMING, D. R. J. (1968).   *Information Theory of Choice Reaction Times*.   London: Academic Press.

RABBITT, P. M. A. (1966).   Errors and error correction in choice-response tasks. *Journal of Experimental Psychology*, **71**, 264–72.

RABBITT, P. M. A. (1968).   Three kinds of error-signalling responses in a serial choice task. *Quarterly Journal of Experimental Psychology*, **20**, 179–88.

RABBITT, P. M. A. and VYAS, S. M. (1970).   An elementary preliminary taxonomy for some errors in laboratory choice RT tasks.   In SANDERS, A. F. (Ed.) *Attention and Performance III*, *Acta Psychologica*, **33**, 56–76.

SCHOUTEN, J. F. and BEKKER, J. A. M. (1967).   Reaction time and accuracy.   *Acta Psychologica*, **27**, 143–53.

STONE, M. (1960).   Models for choice reaction time. *Psychometrika*, **25**, 251–60.

WELFORD, A. T. (1960).   The measurement of sensory-motor performance: survey and reappraisal of twelve years' progress. *Ergonomics*, **3**, 189–229.

CONDITIONAL PROBABILITY AND COMPONENTS OF RT IN THE VARIABLE FOREPERIOD EXPERIMENT

IVOR STILITZ*

University College, London

If the passage of time during the foreperiod in a variable foreperiod experiment is marked by a series of tones, RT decreases with the conditional probability of stimulus occurrence. RTs at short foreperiods, however, are rather slower than would be expected on the basis of a simple conditional probability effect. It is suggested that this is attributable to an independent "initial slow reaction" effect, and it is shown that the degree of this effect is influenced by the duration of the prior foreperiod. The results are related to those of variable foreperiod a-reactions in which no conditional probability effect has been found, and it is argued that in a marked reaction of the kind described above, the initial slow reaction effect behaves like an a-reaction component of RT. It is suggested that the absence of a conditional probability effect in the a-reaction and its presence in the marked reaction are related to the fact that a different type of sensory process is used to identify the signal in each case.

Introduction

The data obtained from a variable foreperiod experiment are typically represented by a plot of RT against foreperiod: an "RT curve". Many studies have used a minimum foreperiod in the region of $\frac{1}{2}$ or 1 sec and a maximum foreperiod of up to 25 sec. In such studies, the form of the RT curve has often been a negatively accelerated decreasing function. This has led a number of authors to suggest that RT decreases as a function of the subject's expectancy of the signal's arrival. Audley (1963) suggests, in addition, that a natural measure of expectancy would be the conditional probability (CP) of stimulus occurrence. In the present context, CP refers to the probability that a signal will occur at a given point in time, given that it has not already occurred. A simple test of the CP hypothesis is difficult to make because the effects of the subject's uncertainty about the possible times of stimulus occurrence are likely to distort any underlying RT/CP relationship. This led Thomas (1967) to develop a number of CP models which take subjective time uncertainty effects into account, and these predict RT curves of a form similar to those usually obtained. Other types of hypothesis, however, could also predict such curves, and Stilitz (1970) has suggested that a variety of data would be more easily explained in terms of a "prepared reflex" concept similar to that discussed by Exner and Cattell (Woodworth, 1938).

The idea behind the present experiment was that if CP affects speed of response in the a-reaction, it might also be expected to affect speed of response in other types

---

*Now at The Hebrew University of Jerusalem.

of reaction in which the foreperiod is variable. It will be shown below that by departing from the standard a-reaction paradigm, it becomes possible to remove the differential effects of time uncertainty on foreperiods of different duration. Hence, it should also become possible to demonstrate a CP effect in a relatively straightforward fashion.

Method

The key aspects of the method used were as follows. (a) A brief tone burst occurred every 2 sec during a trial. These tones marked the passage of time. (b) The subject had to respond to a tone of another intensity which occurred 1 sec after one or other of the "markers". Thus the time uncertainty associated with any signal requiring a response was 1 sec, whatever the interval between that signal and the beginning of the trial. The possible times of occurrence of signal and marker were made different in an attempt to maintain maximum comparability with ordinary a-reactions, i.e. since, in theory, it would be possible for the subject to use a detection process to identify the signal, the condition could be conceived as a "marked a-reaction".

No markers occurred after the response signal, but at least two markers preceded it; i.e. there was dummy marker prior to the first relevant marker. It was felt to be important that the subject have a clear idea of the temporal properties of the situation, and it was for this reason that the subject was given at least one experience of the 2 sec between-marker interval on each trial. On the trials on which a response signal was presented, it was preceded by either 2, 3 or 4 markers. These three contingencies will be called the "1 sec", "3 sec" and "5 sec" foreperiod (FP) respectively, since the durations in question define the interval between the first relevant marker and the response signal. A catch trial consisted of five markers.

In order to explore the possible role of stimulus intensity, one group of subjects were tested on a loud marker tone/soft signal tone (LS) condition, whilst a second group were tested on a soft marker tone/loud signal tone (SL) condition. Each subject was tested on two sub-conditions which differed in terms of the FP probability density function used. The relative frequencies of the FPs in each sub-condition, and the associated values of CP and 1/CP are shown in Table 1. The main reasons for choosing the particular foreperiod probability density functions shown in the Table are as follows:

TABLE I

	Sub-condition A				Sub-condition B			
FP (sec)	1	3	5	Catch trials	1	3	5	Catch trials
Relative frequency	0·5	0·2	0·2	0·1	0·2	0·2	0·5	0·1
CP	0·5	0·4	0·67		0·2	0·25	0·83	
1/CP	2·0	2·5	1·5	—	5·0	4·0	1·2	—

(1) As noted in the Introduction, it has generally been found that RT decreases with foreperiod duration. This could represent some kind of effect whereby RT decreases mainly as a function of time, per se, rather than as a function of CP. It could equally reflect the fact that RT decreases with the cumulative density function, which is necessarily monotonic on FP duration. For these reasons, it was decided to use at least one condition in which CP was non-monotonic on foreperiod. (Both Thomas (1970) and Requin and Granjon (1969) aimed specifically at assessing the possible role of CP. But, in both cases, they only used conditions in which CP was monotonic on foreperiod duration. Their results, which

were broadly compatible with a CP hypothesis, were also compatible with other hypotheses of the kind just mentioned.)

(2) Some preliminary studies suggested that there might be some tendency for RT to decrease with FP independently of any possible CP effect. If this were so, the use of two distributions would make it easier to assess the contributions of the two effects in question.

(3) The use of two distributions enabled each foreperiod to be presented reasonably frequently whilst giving a large spread of CP values (0·2 to 0·83 compared with a maximum possible range of 0·0 to 1·0). The relative frequencies used in the two sub-conditions also provided a control for a "relative frequency" effect of the kind suggested by Mowrer (1940).

The sequences of FPs were random and examination showed that they were free of first order sequential dependencies. The subject's response was to release a morse key. The brief onset of a light indicated that a trial had begun and that the subject must depress the key. The first marker occurred 1 sec later. There was a 3 sec delay between the occurrence of the response signal (or, in the case of a catch trial, the last marker) and the beginning of the next trial.

Subjects

The subjects were 8 male students.

Apparatus

The subject sat alone at a table in a darkened room. Trial initiation was indicated by a red neon light situated at eye level about 45 cm in front of the subject. The light remained on for approximately 0·5 sec. A sine wave generator produced 750 Hz tones at either 79 dB SPL (in the case of the "loud" signal) or 64 dB (in the case of the "soft" signal). The stimuli were delivered to the subject through a pair of headphones via an electronic gate which produced tones of 27–30 msec duration. The sequence of events defining the trial was controlled by a process timer. The timer fired every second, moving-on a paper tape reader. The occurrence of the tones and the light, and resetting of the Decatron timer which recorded RTs to the nearest msec, were controlled by the tape via a relay switching circuit. The experimenter was also able to speak to the subject through the head-phones.

Procedure

Four of the subjects were tested on the LS condition and four on the SL condition. The subjects were tested in five sessions on consecutive days. Sub-conditions A and B were used alternately, the pattern being balanced over subjects within each condition. The Day 1 session was regarded as practice, half of it being run on A and half on B. The schedule of conditions is given in Table II.

TABLE II

Subject	Stimuli	Day 1	Days 2 and 4	Days 3 and 5
S1 S5	S/L	A/B	A	B
S2 S6	S/L	B/A	B	A
S3 S7	L/S	A/B	A	B
S4 S8	L/S	B/A	B	A

In sub-condition A there were nine blocks of 20 trials, 180 trials in all, per session. In sub-condition B, there were eight blocks of 20 trials, giving 160 trials per session. (Slightly fewer trials were given in sub-condition B in order to keep session duration down to about 40 min.) On Day 1, half the number of trials normally used was given in each sub-condition. As well as the first session, the first block in each session and the first trial in each block were omitted from the analysis. The structure of the task was explained to the subject and he was told that FPs would vary from trial to trial. But he was not given any information about the relative frequencies associated with each foreperiod. He was instructed to respond as quickly as possible on each trial and to avoid anticipatory responses. Knowledge of results was given at the end of each block by reading out to the subject the 20 RTs that had just been obtained.

Results

The median RT at each foreperiod was calculated for each subject. The individual data are based on the means of the medians obtained in the different sessions; the grouped data are the means of the individual data. The grouped
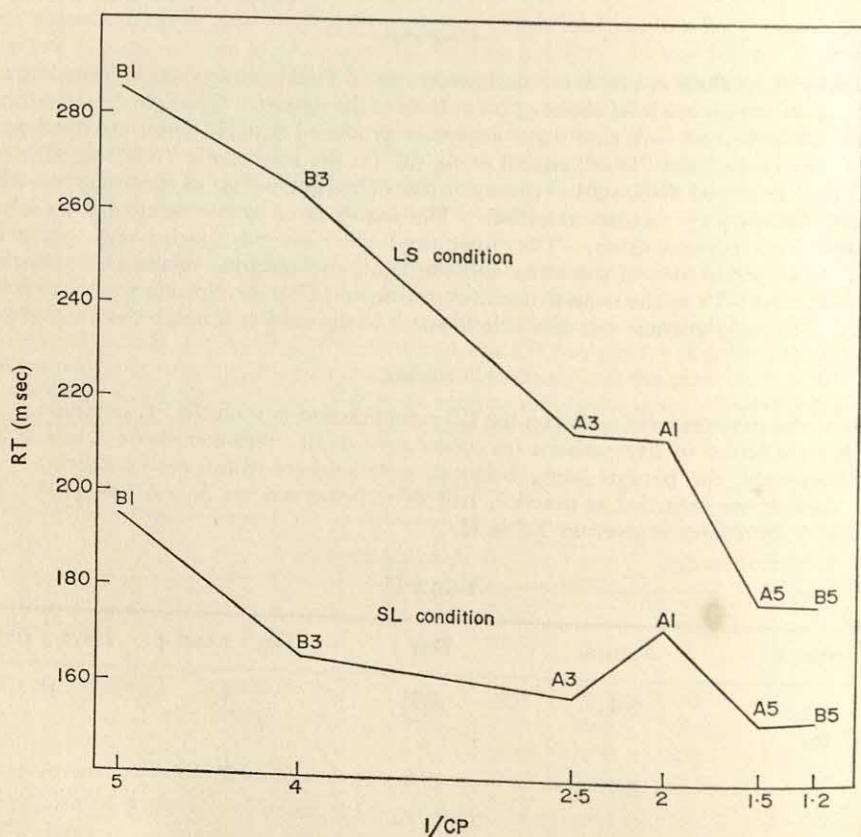

FIGURE 1. RT and CP. The letter next to each point indicates the sub-condition from which the datum was derived; the number indicates foreperiod duration (sec).

data are presented in Figure 1, where RT is plotted against 1/CP. (The use of 1/CP is arbitrary—it gives an approximately linear relationship with RT. It is

not implausible, however, that RT might be reciprocally related to expectancy.)
In the LS condition, RT decreases markedly with CP, and a trend in the same
direction is apparent in the SL data, though the effect is much diminished.   It can
be seen from Table III that the pattern of results shown in the grouped data is
representative of that shown by the individual subjects. (The mean decrease
in RT over the range of CP values is significantly greater in the LS condition;
$t = 3 \cdot 17$, $\nu = 6$, $P < 0 \cdot 02$.)

The relationship between RT and CP would more nearly approximate a smooth
monotonic relationship if RT at the 1 sec foreperiod in the two SL sub-conditions
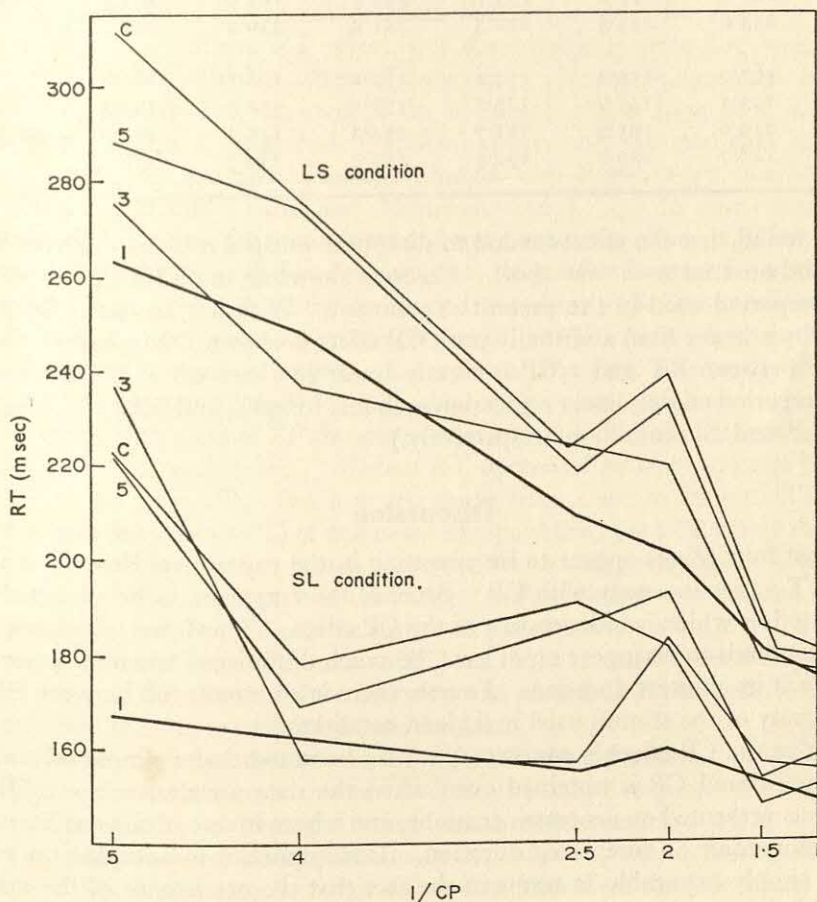
FIGURE 2.   RT and CP: prior foreperiod effects.   The number next to the first point on each
curve shows the prior foreperiod duration (sec) for the curve in question.   C indicates that the
preceding trial was a catch trial.

and in sub-condition (A) of the LS condition were somewhat faster.   Note that
this is true for any simple monotonic scaling of CP.   Stilitz (1970) found what may
be interpreted as a similar "initial slow reaction effect" in ordinary a-reactions.

TABLE III

	Sub-condition A			Sub-condition B			
FP 1	1	3	5	1	3	5	
1/CP	2·0	2·5	5·0	5·0	4·0	1·2	
S3	226·5	223·0	198·0	320·0	277·5	187·2	
S4	188·7	217·7	157·7	272·5	275·0	168·5	LS
S7	180·0	177·2	153·7	196·7	185·0	142·5	condition
S8	253·0	235·5	202·5	351·2	316·0	212·0	
S1	167·0	159·2	154·2	187·2	164·5	161·7	
S2	153·5	146·7	142·7	177·2	151·5	143·5	SL
S5	219·0	195·0	181·7	254·2	214·5	189·2	condition
S6	148·7	129·0	130·5	159·7	130·7	118·5	

He also found that the effect tended to disappear on trial n if the duration of the foreperiod on trial $n-1$ was short. Figure 2 shows separate RT curves for each prior foreperiod used in the present experiment. With a 1 sec prior foreperiod (shown by a heavy line) a virtually pure CP effect is shown. Moreover, the relationship between RT and 1/CP is clearly linear in these cases. (For the 1 sec prior foreperiod curves, linear regression accounts for 98% and 86% of the variance in the LS and SL conditions respectively.)

Discussion

At least four effects appear to be operating in the experiment described above. First, RT varies inversely with CP. Second, there appears to be an initial slow reaction effect which is independent of the CP effect. Third, the initial slow reaction effect tends to disappear or, at least, be much diminished when the prior foreperiod is at its shortest duration. Fourth, there is an interaction between CP and the intensity of the stimuli used in a given condition.

As far as the CP effect is concerned, it may be noted that a simple relationship between RT and CP is obtained even when the data are drawn from different conditions presented on separate occasions, and where in one of the conditions CP is non-monotonic on foreperiod duration. It is important to note that the results are not simply explicable in terms of the fact that the occurrence of the markers has enabled the subject to treat each marker as the effective beginning of a new trial. (We know that in a fixed foreperiod condition, RT varies with the relative frequency of the signal (Gordon, 1967).) The subject's use of conditional probabilities shows, ipso facto, that he is responding to the properties of a set of foreperiods. So does the presence of the initial slow reaction effect, which would be uninterpretable on the basis of a "fixed foreperiod hypothesis". (In a fixed foreperiod condition, RT is always relatively fast at short foreperiods.)

The idea of conducting the marked experiment came about through an interest

in the possible role of CP in the ordinary a-reaction. But although the CP effect in the present experiment is clear-cut enough, Stilitz (1970) obtained data which positively suggest that there is no CP effect in the a-reaction. (Most of the a-reaction studies in question were conducted in parallel with the marked reaction study reported in this paper.) He used foreperiods which varied between 1 sec and 3 sec, or 1 sec and some other value up to a maximum of 17 sec. Rectangular and various non-rectangular foreperiod density functions were used in a number of experiments. The following aspects of the results are relevant to the present discussion. (1) RT was found to be remarkably uniform over most FPs in most of the conditions studied. (2) This uniformity was only clearly manifest when the individual data were inspected. Some of the subjects did not show the uniform RT effect, but these subjects were not consistent in showing any other particular pattern of behaviour. It may be noted that in almost all the a-reaction studies reported in the literature which have used foreperiod ranges of the orders of magnitude discussed above, only grouped data have been presented (cf., for example, Huston, Shakow and Riggs, 1937, Karlin, 1959; Botwinnick and Brinley, 1962, and Naatanen, 1963). (3) In many conditions, RT at the very shortest foreperiods tended to be relatively slow. (4) The relative slowness of initial FPs tended to decrease if (a) the relative frequencies of the initial FPs was high, and (b) if the prior foreperiod was short. (Cf. Zahn and Rosenthal, 1966.) For the very shortest prior foreperiod, an RT curve that was uniform throughout was sometimes found. Examples of the form of overall RT curves (disregarding prior foreperiod) for individual subjects are as follows: (i) a subject was presented with a set of FP durations which included all the integral multiples of 1 sec between 1 and 17 sec. Median RT decreased by over 50 msec from the 1 sec FP to the 3 sec FP. But over the range from 4 sec to 17 sec, RT stayed within 5 msec (less than $2\frac{1}{2}\%$) of 226 msec, except at the 13 sec FP where the deviation was somewhat larger. It was not possible to detect any trend away from uniformity over the major part of the range. (ii) A subject was presented with the five FPs which were the integral multiples of $\frac{1}{2}$ sec between 1 sec and 3 sec. For the 2 sec, 2·5 sec and 3 sec FPs, median RT was 169, 171 and 170 msec respectively. If there was any CP effect in the conditions just described it was cancelled out by time uncertainty in an extremely precise fashion.

It was concluded that there is no CP effect in the a-reaction. But, note, the behaviour of the a-reaction as described above is in most respects remarkably like that of the initial slow reaction effect found in the marked reaction. The equivalent effects of prior foreperiods is particularly striking; in both types of condition, initial slow reactions disappear or are much diminished (in the case of the marked reaction, in relation to the CP effect) when prior foreperiods take on their shortest value. It seems reasonable to assume, then, that the initial slow reaction effect is an "a-reaction component" of RT in the marked reaction.

Suppose the a-reaction/initial slow reaction effect is attributable to the fact that the subject attempts to maintain a fixed level of motor readiness throughout the duration of a trial but tends to delay the point at which this fixed level is reached if he is not reinforced for its early attainment by the occurrence of short foreperiods. An effect of the kind just described conforms to the traditional notion of the

a-reaction as a "motor response" and would account for the role of prior fore-periods in a fairly natural fashion. (It may be noted that the fact that the prior foreperiod effect occurs in the marked reaction means that this effect is not related to subjective estimates of the location of the minimum foreperiod. It is, moreover, difficult to see why the duration of the prior foreperiod should affect, say, subjective estimates of the conditional probability only for short foreperiods.)

If the initial slow reaction effect is attributable to a motor process, stimulus processing time in the a-reaction must be identical at all FPs. This would happen if the stimulus processing mechanism were pre-programmed at the beginning of a trial so that the response was automatically triggered off by any large increase in sensory input. (A similar notion has recently been discussed by John (1966).)

McGill (1963) and Grice (1968) discuss a latency mechanism in which a response is made when a count of neural impulses (evoked by the onset of a signal) reaches some criterion level. In terms of such a mechanism, the a-reaction process may be interpreted as one in which a liberal fixed criterion is pre-programmed at the beginning of the trial. Now, in the marked reaction, the subject cannot use the type of "automatic" response described above; he would otherwise respond to the markers. In order to identify the signal, he can still briefly bring a detection mechanism into play after the occurrence of a marker, since another marker cannot occur for 2 sec. But the detection mechanism must be continually taken in and out of operation or, at least, "re-set". In such a situation there would be no advantage in using the relatively simple strategy of maintaining a fixed response criterion; on the contrary, it would be expected that the subject would tend to take signal probability into account in fixing his response criterion. If this were so, McGill's model would predict a probability/signal intensity interaction of the kind obtained in the present experiment. (The probability/intensity interaction might be complicated by the interactive effects of the marker and signal intensities, though it is not clear in detail how this would happen. Post-stimulatory masking effects do not extend over 1 sec (cf., for example, Ward, 1963), and the type of warning-signal intensity effect discussed by Kohfeld (1969a, b) suggests that marker intensity and signal probability would have an additive effect on the subject's criterion.)

Given that the subject cannot use an automatic response in the marked reaction, it might be simpler for him not to use detection at all in order to identify the signal. Rather, he could rely on a discrimination mechanism throughout the trial in order to determine whether each stimulus is either a marker or a response signal. In other words, although the marked reaction was originally conceived as a "marked a-reaction", it may have been, in effect, a "marked c-reaction". In this case, too, the subject must "begin again" after each discrimination, and the varying signal probabilities might be expected to be influential. The basis of this discrimination process could be an assessment of the rate of arrival of neural impulses, and a model of the type required might be developed mathematically in a variety of ways. For present purposes, it is sufficient to show that a conditions/probability interaction in the required direction is possible. This can be done by supposing that stable information about the stimulus presented is obtained more quickly for a louder stimulus and hence criterion shifts will have a greater effect on the more variable

information arising from a weaker stimulus. More specifically, suppose the inter-arrival time of impulses is random and that the mean rate of arrival is a function of stimulus intensity. Suppose further that the subject makes a continuous assessment of the mean inter-arrival time and of the variance about this mean. He makes an appropriate decision, given the mean, when the variance is less than some criterion level. The criterion is varied according to the probability of the signal. Now a given number of inter-arrival times will have a smaller variance in the case of a relatively loud signal. It will, moreover, have taken less time to accumulate this number of inter-arrival times. Since relevant information is accumulating more quickly in the case of a loud signal, a given shift in criterion will have a lesser effect.

The main points in the argument above are as follows : (a) The initial slow reaction effect in the marked reaction behaves like an a-reaction component of RT. (b) No CP effect would be expected in the a-reaction if the subject sets himself up to respond to any gross increase in sensory input. (c) He cannot set himself up in this manner in the marked reaction, and signal probability effects would then be expected to come into play.

The particular explanations suggested above have been put forward in the spirit of being fairly natural given the nature of the data obtained; clearly, alternative explanations are possible on a variety of points. Since CP interacts with stimulus intensity, it seems reasonable to suppose that it is affecting a sensory component of RT. But it has not been shown that the initial slow reaction effect does not do so. It might be argued that expectancy is a weighted function of overall CP, experience on immediately preceding trials, and perhaps other variables. This expectancy might then influence RT through the operation of a single mechanism by determining, say, a response criterion. On the other hand, the supposition that the initial slow reaction effect represents a motor component, or at least is attributable to some independent process, enables the differences in results between the a-reaction and marked reaction to be dealt with in a fairly straightforward fashion. The presence of markers does not prevent the subject maintaining, say, a fixed level of motor readiness throughout the duration of a trial, but it does prevent him from being prepared simply to respond to any gross increase in sensory input.

A number of further experiments should clarify the issues in question. First, comparison of a-reaction RT curves, for different signal intensities, would test the hypothesis that these curves should be identical except for their displacement on the ordinate. Second, comparison of a marked c-reaction and the type of marked reaction described in this paper, using SL and LS conditions in both cases, should throw light on the question of what type of signal identification mechanism determined the results obtained in the present case. Third, parametric studies of marked reactions should enable the precise inter-relationship of the CP effect, the initial slow reaction effect, the prior foreperiod effect, and the signal intensity effect to be determined.

References

AUDLEY, R. J. (1963). Decision models in reaction time. In *Psychophysics and the Ideal Observer*, Proceedings of the 17th International Congress of Psychology. Amsterdam: North-Holland Publishing Company.

BOTWINNICK, J. and BRINLEY, J. F. (1962). An analysis of set in relation to reaction time. *Journal of Experimental Psychology*, **63**, 568–74.

GORDON, I. E. (1967). Stimulus probability and simple reaction time. *Nature*, **215**, 195–6.

GRICE, G. R. (1968). Stimulus intensity and response evocation. *Psychological Review*, **75**, 359–73.

HUSTON, P. E., SHAKOW, D. and RIGGS, L. A. (1937). Studies of motor function in schizophrenia: II Reaction time. *Journal of General Psychology*, **16**, 39–82.

JOHN, I. D. (1966). Intensity of non-key stimuli in the Donders' c-type reaction. *Australian Journal of Psychology*, **18**, 148–53.

KARLIN, L. (1959). Reaction time as a function of foreperiod duration and variability. *Journal of Experimental Psychology*, **58**, 185–91.

KOHFELD, D. L. (1969a). Effects of the intensity of auditory and visual ready signals on simple reaction time. *Journal of Experimental Psychology*, **82**, 88–95.

KOHFELD, D. L. (1969b). Effects of ready-signal intensity and intensity of the preceding response signal on simple reaction time. *American Journal of Psychology*, **82**, 104–10.

McGILL, W. J. (1963). Stochastic latency mechanisms. In LUCE, R. D., BUSH, R. R. and GALANTER, E. (Eds), *Handbook of Mathematical Psychology*. New York, London: Wiley.

MOWRER, O. H. (1940). Preparatory set (expectancy)—some methods of measurement, *Psychological Monographs*, **52**, No. 233.

NAATANEN, R. (1963). On the interrelations of foreperiod and reaction time in simple reaction experiments. *Reports from the Psychological Institute, the University of Helsinki*, No. 1.

REQUIN, J. and GRANJON, M. (1969). The effect of conditional probability of the response signal on the simple reaction time. *Acta Psychologica*, **31**, 129–44.

STILITZ, I. B. (1970). Readiness and Simple Reaction Time. Doctoral thesis, University of London.

THOMAS, E. A. C. (1967). Reaction-time studies: the anticipation and interaction of responses. *British Journal of Mathematical and Statistical Psychology*, **20**, 1–29.

THOMAS, E. A. C. (1970). On expectancy and average reaction time. *British Journal of Psychology*, **61**, 33–8.

WARD, W. D. (1963). Auditory fatigue and masking. In JERGER, J. (Ed.), *Modern Developments in Audiology*. London: Academic Press.

WOODWORTH, R. S. (1938). *Experimental Psychology*. New York: Holt.

ZAHN, T. P. and ROSENTHAL, D. (1966). Simple reaction time as a function of the relative frequency of the preparatory interval. *Journal of Experimental Psychology*, **72**, 15–19.

ACUTE ANGLES AND THE POGGENDORFF ILLUSION†

A. W. PRESSEY AND O. SWEENEY‡

Department of Psychology, University of Manitoba

Burns and Pritchard's (1971) explanation of the Poggendorff illusion is criticized. An experiment was designed to determine whether the acute angle plays any role in the perception of the illusion. The results showed that (i) an inducing line which crossed a test-line was highly effective in altering the apparent orientation of the test line, (ii) an inducing line forming an acute angle with a test-line had a small effect in changing the apparent orientation of the test-line, and (iii) an acute angle which formed part of the Poggendorff configuration produced an effect opposite to that predicted by the view that acute angles are perceptually enlarged.

Introduction

In 1968, Chiang presented a neurophysiological theory to explain illusions such as the Poggendorff, Zöllner and Müller-Lyer. Subsequently this theory was severely criticized; yet, in 1971, Burns and Pritchard presented a model which is, in its predictive features, almost identical to Chiang's theory.

Burns and Pritchard contend that, when one views an acute angle, there is a systematic distortion of the cortical image of the sort shown in Figure 1. Points
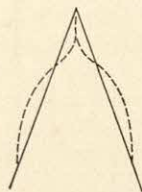
FIGURE 1. Hypothetical distortion of image (dashed lines) produced by an angle (solid lines).

which are close together are displaced toward each other and more distant points (up to several degrees) are displaced away from each other. The result is that acute angles are enlarged subjectively and, according to Burns and Pritchard, this provides a *sufficient* explanation for the Poggendorff illusion. They go on to say that "if a particular illusory percept can be explained as a consequence of distorted sensory input, then its importance in a continuing study of brain functions becomes trivial (p. 615)". Unfortunately, Burns and Pritchard appear to be unaware of the

literature on the Poggendorff illusion and, in view of this, a brief review of some observations seems necessary.

It has been pointed out by Judd (1899) and reiterated by Pressey and den Heyer (1968) that if, in the Poggendorff illusion, all contours except those forming acute angles are removed, the illusion disappears. Restle (1969), in a quantitative study, showed that, not only does the illusion disappear, but the distortion is actually in the reversed direction. His results substantiated an earlier report by Green and Hoyle (1964) in which a small reversed illusion was demonstrated. Clearly, the abolition or reversal of the Poggendorff illusion when acute angles are the only kinds of angles present violates any theory which purports to provide a sufficient explanation on the basis of the presence of acute angles.

A second difficulty for Burns and Pritchard is provided by the target shown in Figure 2(a). Here one vertical line from the Poggendorff illusion is removed
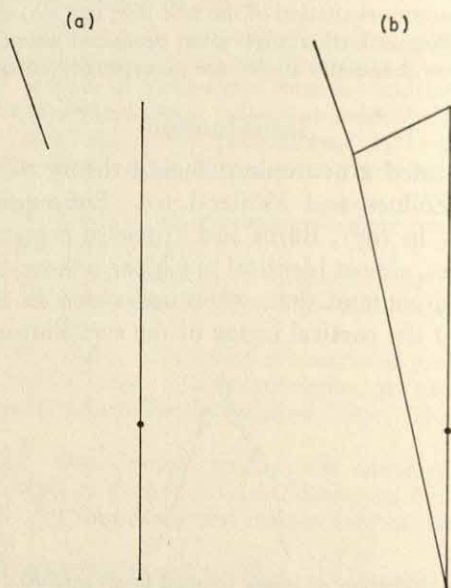
FIGURE 2. (a) A Poggendorff illusion without closed angles; (b) a reversed Poggendorff illusion.

and one oblique is replaced by a dot. This provides a target in which there is no closed acute or obtuse angle. Burns and Pritchard would have to predict that the illusion should disappear. But Kobayashi (1956) and Pressey and Sweeney (1969) found a substantial illusion in the classical direction when a target similar to Figure 2(a) was employed. Moreover, this result cannot be explained on the basis of Weintraub and Virsu's (1971) hypothesis that observers reduce the angular disparity between a line segment and the nearer cardinal axis because, with the targets employed, a reversed illusion would have been expected. Furthermore, it cannot be argued that the target shown in Figure 2(a) is entirely different from the classical illusion since there was a highly significant correlation between the size of the illusion produced in individual subjects by target 2(a) and by the normal Poggendorff figure including a closed acute angle (Pressey and Sweeney, 1969).

Finally, if target 2(a) is modified in the manner shown in Figure 2(b), the illusion *reverses* (Pressey and Dewar, 1970).   It is not at all clear how Burns and Pritchard's model could be extended to explain the reversed illusion.

Thus, the fact that an illusion exists when all closed angles are removed shows that distortion of angles is neither a necessary condition nor a sufficient explanation for the Poggendorff illusion.   Nevertheless, an angle theorist might still argue that distortion of angle, although neither necessary nor sufficient, is still a powerful factor in the actual perception of the Poggendorff configuration.   The purpose of the present study was to determine whether an acute angle plays *any* role in the Poggendorff illusion and, if so, how much it contributes to the distortion.

Method

Materials

Seven different targets, shown in Fig. 3, were employed.   The basic target (3(a)) consisted of black lines drawn on 28 × 21·7 cm sheets of white paper.   The vertical line
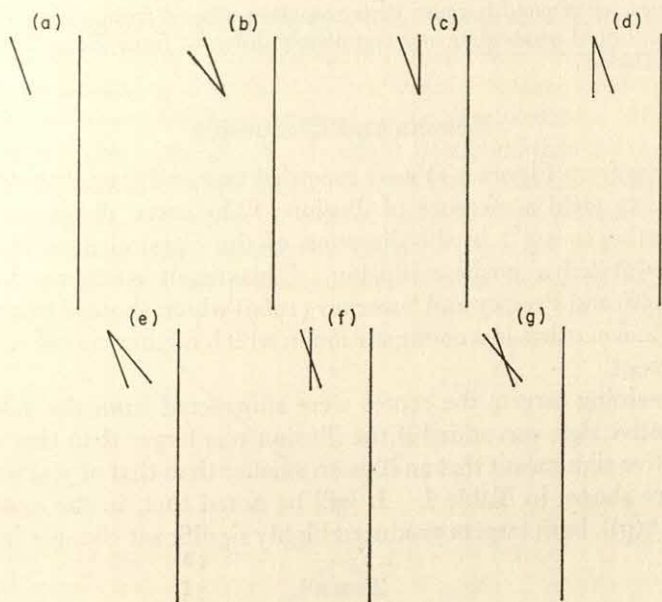
FIGURE 3.   Targets employed in the experiment.

was 17·8 cm long and the oblique line (test-line) on the left was 3·7 cm long.   The angle formed by the test-line was 20° from the vertical axis and, if extended, it joined the vertical line at a point 5·1 cm from the bottom of that line.   The width of all lines was approximately 0·5 mm.   The remaining targets were constructed by adding a line (with a distinct gap at one extreme) to the test-line to form acute angles as shown in Figure 3.   The targets were reproduced by printing, and booklets containing all seven targets, ordered in random fashion, were constructed.   Each target was preceded by a heavy blue sheet of paper to prevent the target beneath from being visible.

Procedure

Subjects were tested in groups of 8–12. They were seated at tables which were situated in such a manner that subjects could not view the responses of others. A white dot was located on the table on which the preferred hand rested between trials. The subject was shown a dummy target which was similar to Fig. 3(c) but with the test-line forming an angle of 70° with the vertical axis. He was told to look at the dummy card and to notice that one line was solid and that the other had a gap at the end. He was also told that his task was to project visually the *solid* line and to draw a dot on the long vertical line at a point where the solid line would intersect if it were extended. He was asked to keep his hand on the white spot and remove it only after the signal "Now". After the response, he was to turn the white paper, replace his hand on the dot, and wait for the next signal. Responses were made on the dummy card and these were collected and checked by the experimenter to ensure that subjects understood the instructions.

The experimental trials began with the experimenter giving the signal "Turn". Subjects turned the blue sheet and inspected the target for 10 sec. The experimenter then signalled "Now", the subject made his response and, immediately after, turned the white sheet. Trials continued until responses were made on all seven targets.

Subjects

The subjects were 66 students enrolled in introductory psychology at the University of Manitoba. Only those with good vision (with or without glasses) were asked to participate. Two subjects were eliminated because their responses ranged from 1 mm to 125 mm and because their pattern of responding was completely different from the pattern exhibited by the remaining subjects.

Results and Discussion

The responses from Figure 3(a) were recorded first and the value was subtracted from the POE to yield a measure of illusion. The mean illusion was 17·9 mm (angular distortion = 4·5°) in the direction of the classical illusion. Of the 64 subjects 56 exhibited a positive illusion. This result confirms the studies of Kobayashi (1956) and Pressey and Sweeney (1969) which showed that a substantial Poggendorff illusion exists in a configuration in which neither closed acute or obtuse angles are present.

For the remaining targets the scores were subtracted from the value on target 3(a) and a positive sign was added if the illusion was larger than that exhibited on 3(a). A negative sign meant that an illusion smaller than that of 3(a) was exhibited. The results are shown in Table I. It will be noted that, in the case of crossing lines (3(f) and 3(g)), both targets produced highly significant changes in illusion in a

TABLE I

Change in illusion (mm) produced by added angles

Target	(b)	(c)	(d)	(e)	(f)	(g)
Distortion predicted by angle theory	−	+	+	−	+	−
Mean	1·32	−9·17	1·48	−4·74	4·25	−8·39
t	0·74	4·88	1·10	3·84	3·02	4·43
P	N.S.	<0·01	N.S.	<0·01	<0·01	<0·01

direction consistent with the view that acute angles are perceptually enlarged. However, of the four cases in which a single acute angle was employed, only the results of target 3(e) are significant in the predicted direction; those for target 3(c) are significant in the *reversed* direction and this is precisely the target in which the acute angle forms part of the Poggendorff configuration.

What can be concluded from these results? First, it seems that acute angles can distort the contours which form the arms of the angle and, in this sense, small angles appear to grow. It is also clear that, when the acute angle forms part of the Poggendorff configuration, it does not enhance, but rather decreases, the magnitude of illusion. Thus, it is exceedingly difficult to defend the view that, in the Poggendorff illusion, acute angles cause the distortion or even play any role in enhancing the illusion. Yet, in other kinds of illusions, such as the Hering figure, acute angles could play a role despite the fact that *acute angles are unnecessary* for this illusion to occur. Coren (1970) replaced the solid parallel lines with dots and found that a substantial illusion was still exhibited. Perhaps the best conclusion that can be drawn is that acute angles can, in certain cases, enhance geometric illusions, but that, in most cases, they contribute little to the perceived distortion.

The results of this study also show that inducing lines which actually cross the test-line are powerful factors in changing the apparent orientation of the test-line. In a sense, there is nothing new about such a finding because Gibson's (1937) early studies on after-effects of tilt demonstrated such a phenomenon. What is troublesome is the explanation of the effect. It might be argued that the models provided by Chiang (1968) and by Burns and Pritchard (1971) are sufficient. According to both of these models, the sensory input for an acute angle is modified in a manner approximately as shown by the dashed line in Figure 1. But, if the model is pursued for crossing lines, the predicted percept is as shown in Figure 4. How-
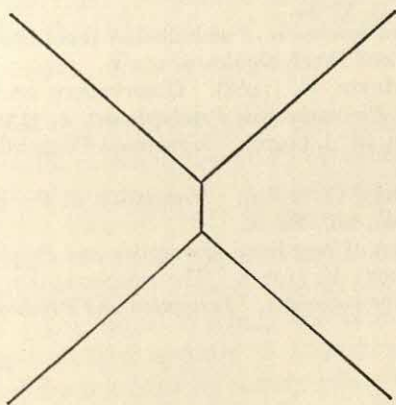

FIGURE 4. Predicted percept produced by crossing lines according to model shown in Figure 1.

ever, crossing lines just do not look like that; we perceive two straight lines crossing each other. In order for the two models to predict the actual percept they would have to contend that the organism takes the "line of best fit" through the pattern

shown in Figure 4. But when (and if) it does so, we are left with the original figure consisting of crossed lines with no illusion whatsoever.

An alternative point of view is that an inducing line which crosses a test-line alters the orientation of the entire line so that it appears to rotate with the crossing point being the centre of rotation. This rotation could be due to lateral inhibition between orientation detectors as suggested by Blakemore, Carpenter and Georgeson (1970) or perhaps it could be due to the manner in which information is processed along the lines suggested by Pressey (1971) in his theory of geometric illusions. It is true that Burns and Pritchard found "no evidence that the orientation of a single test line preferred by cortical neurons was different from that preferred by the same cell when excited by an angle pattern, one arm of which was the original test line" (p. 599), but one wonders whether they would have reached the same conclusion if a crossing line, rather than an angle, had been used as the inducing figure.

References

BLAKEMORE, C., CARPENTER, R. H. S. and GEORGESON, M. A. (1970). Lateral inhibition between orientation detectors in the human visual system. *Nature*, 228, 37–9.

BURNS, B. D. and PRITCHARD, R. (1971). Geometrical illusions and the response of neurones in the cat's visual cortex to angle patterns. *Journal of Physiology*, 213, 599–616.

CHIANG, C. (1968). A new theory to explain geometrical illusions produced by crossing lines. *Perception and Psychophysics*, 3, 174–6.

COREN, S. (1970). Lateral inhibition and the Wundt-Hering illusion. *Psychonomic Science*, 18, 341–2.

GIBSON, J. J. (1937). Adaptation, after-effect, and contrast in the perception of tilted lines. II. Simultaneous contrast and the areal restriction of the after-effect. *Journal of Experimental Psychology*, 20, 553–69.

GREEN, R. T. and HOYLE, E. M. (1964). The influence of spatial orientation on the Poggendorff illusion. *Acta Psychologica*, 22, 348–66.

JUDD, C. H. (1899). A study of geometrical illusions. *Psychological Review*, 6, 241–61.

KOBAYASHI, T. (1956). Analytical study of displacement in visual perception (I). *Japanese Psychological Research*, 3, 37–44.

PRESSEY, A. W. (1971). An extension of assimilation theory to illusions of size, area and direction. *Perception and Psychophysics*, 9, 172–6.

PRESSEY, A. W. and DEN HEYER, K. (1968). Observations on Chiang's "new" theory of geometrical illusions. *Perception and Psychophysics*, 4, 313–14.

PRESSEY, A. W. and DEWAR, M. J. (1970). A reversed Poggendorff illusion. *Psychonomic Science*, 20, 227–8.

PRESSEY, A. W. and SWEENEY, O. (1969). A variation of the Poggendorff illusion. *Perceptual and Motor Skills*, 28, 883–6.

RESTLE, F. (1969). Illusions of bent line. *Perception and Psychophysics*, 5, 273–4.

WEINTRAUB, D. J. and VIRSU, V. (1971). The misperception of angles: Estimating the vertex of converging line segments. *Perception and Psychophysics*, 9, 5–8.

THE PSYCHOLOGICAL REFRACTORY PERIOD AND VERTEX EVOKED POTENTIALS

JOHN BODDY

Psychology Department, University of Manchester

The averaged sensory evoked potential (EP) was recorded from the scalp (vertex to mastoid) in a psychological refractory period experiment in which 12 young adults participated. Reaction times (RTs) were measured to either both or only the second of pairs of stimuli, in different trial blocks, with inter-stimulus intervals (ISIs) of 100, 200, 300 and 400 msec occurring in random sequence. EPs were recorded at each ISI. No latency changes could be found in the prominent non-specific components (P1–N1–P2) of the EP to stimulus 2 even at ISIs where the RT was substantially delayed. Thus the notions that the RT2 delay is due to occupation of a single channel central processor by S1 and that non-specific EP components reflect the time course of information processing in underlying neural tissue, do not lend each other mutual support. Furthermore, as profound amplitude refractoriness in components P1–N1 and N1–P2 persisted at ISIs where RT was as fast or faster than simple RT, there appears to be a dissociation between "psychological refractoriness" and "physiological refractoriness". The implications of these results are discussed.

Introduction

The notion of a psychological refractory period (PRP) to explain the limitations on human information processing in a rapid temporal sequence, was introduced by Telford (1931). The associated phenomena generated sustained interest in the 1950s and 1960s following the elaboration of the linked concept of central intermittency by Craik (1947, 1948). Bertelson (1966) and Smith (1967a) have reviewed research in this area.

The usual paradigm for the investigation of the phenomena subsumed under the heading PRP is the serial reaction time (RT) experiment. The usual finding is that the RT to the second of two closely succeeding stimuli lengthens progressively as the inter-stimulus interval (ISI) is shortened under about 300 msec. Amongst conflicting attempts to explain the delay in the second response the notion that it is a function of the operation of a single channel central processor has proved most difficult to refute. Any signal reaching the central mechanism before a previous signal has been dealt with appears to be held up until the processor is cleared and has recovered from a brief refractory phase following any operation. As the leading protagonist of this theory Davis (1956, 1957, 1958, 1962, 1965) has shown, in a neatly interlocking series of experiments, that there is an irreducible minimum delay in a response to the second of two closely succeeding stimuli, occurring even when the first stimulus demands no response, which can only be attributed to a single channel central mechanism. It cannot be explained

in terms of delay in the sensory or motor tracts or the effects of temporal uncertainty (Davis, 1965).

The term "psychological refractory period" was coined because the phenomena observed appeared analogous to physiological refractoriness. Recent developments in electrophysiology have made possible the gross recording of neural correlates of some psychological processes. A prominent example of such neural correlates is the averaged evoked potential recorded from the scalp, which is a concomitant of the registration of a transient stimulus. It therefore seems appropriate at this stage to assess whether "psychological refractoriness" exhibited in delayed behavioural responses and "physiological refractoriness" displayed by cortical evoked potentials are related.

The nature and significance of evoked potentials (EPs) are far from clear, despite the enormous volume of work on them in the 1960s, heralded by the introduction of averaging techniques (see Goldstein, 1961, for principles) for extracting these signals from noisy backgrounds. Evoked potentials are recordable all over the scalp and comprise a complex sequence of waves lasting 400 msec or more. Bergamini and Bergamasco (1967) have reviewed work on EP morphology. The accumulation of evidence (reviewed by Goff, 1969) indicates that early and late components of the evoked potential are mediated by different independent systems. Components occurring in about the first 100 msec in visual EPs and in the first 65 msec in auditory EPs appear to be mediated by the classical projection pathways, whilst later components are mediated by a diffusely projecting system with its source in the non-specific thalamic nuclei.

There is some debate as to whether the early components of the EP recorded from the scalp are myogenic rather than neurogenic in origin (Bickford, Galbraith and Jacobson, 1963; Bickford, Jacobson and Cody, 1964; Davis *et al.*, 1964), although strong evidence can be marshalled against the assertion (Katzman, 1964; Domino and Corssen, 1964; Domino, Matsuoka, Waltz and Cooper, 1964; Vaughan, 1966).

Whilst the primary segment of the EP is recorded with its greatest amplitude over the sensory projection area corresponding to the stimulus used, the non-specific secondary components tend to be recorded with their greatest amplitude at the vertex, presumably representing spatial averaging over a large area of the cortex. It is this latter portion of the EP which generates the greatest interest for behavioural experimenters as it appears to be systematically modifiable across a variety of conditions, and there is some evidence that it reflects "information processing" occurring in the neural tissue generating the potential (John, 1967). For instance, differentiation between geometrical forms can be seen in the morphology of the EP (John, Herrington and Sutton, 1967). Goff (1969) suggests an intimate relation between the "late potentials" and the neural mechanisms of consciousness, because the late potentials are depressed by anaesthetics or sleep and enhanced when evocative stimuli are the focus of attention. The late potentials referred to exclude a very late negative component (of 260–300 msec latency in the auditory EP) which is augmented during "inattention" (Wilkinson and Morlock, 1967) and sleep (Weitzman and Kremen, 1965; Ornitz, Ritvo, Carr, La Franchi and Walter, 1967). The primary segment of the EP seems a poor can-

didate for a manifestation of neural mechanisms mediating "conscious awareness" or "information processing" as Goff cites evidence that it is unchanged when both are clearly absent, as in sleep or anaesthesia.

A number of researchers have found correlations between augmented non-specific EP components and enhanced performance, attention and selective attention (e.g. Donchin and Cohen, 1967; Haider, 1967; Wilkinson, 1967) and their findings are not inconsistent with the information processing hypothesis.

A psychological theory has been outlined which attributes the prolonged RT2 in serial RT experiments to a delay before processing the second signal in a single channel system. Furthermore, it is hypothesized that evoked potentials reflect the processing of signals reaching the brain. The joint correctness of these two hypotheses would predict that the EP to the second of two closely succeeding stimuli would appear with a greater latency, whose delay beyond the normal value would reflect the delay before the appearance of the behavioural response. Unfortunately, none of the available evidence fits this formulation. No latency or amplitude refractoriness is apparent in the primary response beyond ISIs of about 50 msec (Ciganek, 1964). Davis, Mast, Yoshie and Zerlin (1966) could find no latency shifts in prominent non-specific components of the EP (the "V" potential) to the second of two closely succeeding stimuli. Although no latency shifts were found, Davis et al. (1966) found that, for paired stimuli, there was amplitude refractoriness persisting for up to 10 sec, which is way beyond the ISI where psychological refractoriness turns to facilitation. However, in the experiments by both Ciganek (1964) and Davis et al. (1966) the subjects were passive observers of the evocative stimulus, so that it is important to repeat them with stimuli having imperative significance. It was decided to mount an exploratory serial RT experiment to investigate the fate of the most prominent non-specific components of the EP to stimuli eliciting behavioural responses which were delayed.

Method

Apparatus

Averaging was accomplished using a Data Laboratories Biomac four-channel digital averaging computer receiving an EEG signal from an AEI amplifier. Permanent records of EPs were obtained by outputting stored data to a polyrecorder with a gantry mounted pen for linear write-out. Write-out speeds and sensitivity were selected such that linear measurements on the chart yielded latency measurements with an accuracy ±2·5 msec and amplitude measurements with an accuracy of ±0·4 μV.

A unit constructed of Hawker Siddeley logic modules delivered pairs of stimuli with four possible ISIs of 100, 200, 300 and 400 msec, and four possible inter-pair intervals (IPI) between 4·5 and 7 sec, each category of interval occurring in a random sequence. The device switched the EEG signal between the four channels of the Biomac according to the ISI value selected and delivered a trigger pulse for the Biomac sweep upon the occurrence of each first stimulus (S1). Visual stimuli were 50 msec low intensity green flashes from a Kymore binary indicator. Auditory stimuli were clicks of about 75 dB delivered through headphones suspended above the subjects, rather than worn, as they gave rise to electrical interference. Subjects responded by pressing Bulgin S 506 micro-switches with their forefingers; left hand for S1, right hand for S2. Recycling of the apparatus was triggered by the subjects' response to the second stimulus (S2). The apparatus was used for measuring simple RT with the same temporal pattern (an approximate range of 4·5–7 sec between

each single stimulus, by switching off S2 and operating with experimenter triggered recycling.

Response latencies were measured either by delivering trigger pulses to external Venner stopclocks with automatic reset, or by a crystal clock within the apparatus. In the latter instance data were encoded onto punched tape for computer analysis which yielded means, medians and S.D.s of RT1s and RT2s at each ISI, within each trial block.

Experimental design

Pairs of stimuli were delivered at the four ISIs of 100, 200, 300 and 400 msec. A visual stimulus always occurred first (S1) and an auditory stimulus second (S2). As the auditory stimulus produced higher amplitude EPs than the visual stimulus, attenuated EP2s embedded in an overlapping EP1 were easier to detect when S2 was auditory. The range of ISIs ensured that RTs to S2 would range between being substantially delayed at the shortest ISI and normal or even "supernormal" at the longest. Four stimuli-response paradigms were used in the experiment:

(1) Presentation of 16 visual stimuli followed by 16 auditory stimuli to a passive subject instructed only to inwardly register the occurrence of each stimulus.

(2) 16 simple visual RTs and then 16 simple auditory RTs with knowledge of results. Visual RTs left handed and auditory RTs right handed.

(3) 64 visual-auditory stimulus pairs with response to S2 only.

(4) 64 visual-auditory stimulus pairs with responses to both stimuli.

In the single stimulus situations averaged EPs were based on the sequences of 16 stimuli. In the paired stimulus situation the sequences of 64 pairs yielded EPs for each of the four ISIs, based on an average of 16 stimuli each. When the number of samples contributing to an average deviated from the mean of 16 an appropriate correction factor was applied in amplitude calculations.

The 16 stimulus presentations upon which averaged EPs were based is quite small relative to the numbers used in other studies (e.g. Wilkinson and Morlock (1967) used 50) which sometimes number in the thousands. However, Walter (1964) notes that, in averaging, the gain in signal to noise ratio is proportional to the square root of the number of observations made, for a time coherent signal, so that there is a strict law of diminishing returns for increasing the number of samples. Walter (1964) has successfully averaged over 12 trials. Furthermore, a most important point is that relevant data may be obscured or distorted in averages based on large numbers of samples because, as both Brazier (1964) and Walter (1964) have observed, the likelihood of condition constancy in the brain diminishes rapidly with time and the repetition of stimuli. The experimental findings of Ritter, Vaughan and Costa (1968) and Levonian (1966) lend support to this assumption. Another consideration was that any substantial increase in the number of averaging samples would have led to the experiment being excessively long for the voluntary subjects used, resulting in excessive contamination of the results by boredom and inattention, which would have outweighed any advantages accruing from large sample numbers.

The four stimulus-response paradigms were administered to the 12 different subjects in four different orders such that a latin square was replicated three times in the experiment.

Subjects

Six male and six female subjects were used, all young adults, mean age 23·7 years, S.D. 3·4, range 18–30. 11 subjects were volunteers from amongst students and teachers of psychology and one was a primary teacher.

Procedure

EEG recording was from silver/silver chloride cup electrodes filled with salt jelly and attached with colloidion to the vertex and right mastoid process, the latter electrode being the reference. A one-second time constant was used to avoid attenuating some of the slower EP components and an H.F loss control was set to attenuate activity above 15 cycles, as this tended to include artifacts such as mains hum and muscle activity. The averaging sweep

time was 1·28 sec, so that the EPs to both S1 and S2 were fully recorded at all ISIs. The experiment was conducted in a quiet room illuminated by daylight or artificial light. Subjects reclined on a comfortable couch and were encouraged to relax and avoid gross movements of the bodily and facial musculature. They were instructed to keep their eyes open throughout the experiment, maintain fixation on the visual stimulus source and avoid blinking following the stimuli. The latter instructions were given to try and eliminate eye movement artifact. The visual stimulus source was gantry mounted and adjusted so that it was about 10 in. above the subjects' eyes.

Care was taken to establish good rapport with subjects and all procedures were fully explained to allay any fears and doubts. Practice/warm-up sessions were given to each subject on each task; 10 plus trials on simple RTs and 20 plus on paired stimuli, until a reasonably constant level of performance was achieved. Subjects were instructed to respond as fast as possible and given an immediate knowledge of RTs as an incentive to maximum performance. Subjects were given prior knowledge of the values of the ISIs separating paired stimuli and told that each had an equal probability of occurrence in a random sequence. When responses were required to S2 only, subjects were told to attend to the first stimulus, but avoid responding in anticipation of the second. When responses were required to both stimuli, subjects were instructed to give both equal priority whilst giving a separate response to each stimulus.

Results

The pattern of RTs obtained in this experiment replicated in essentials those obtained in other PRP experiments of this type (e.g. Davis, 1958), although the subjects were much less practised. However, there was a considerable amount of interesting inter-individual variation not seen in the work of Davis as he only used two subjects.
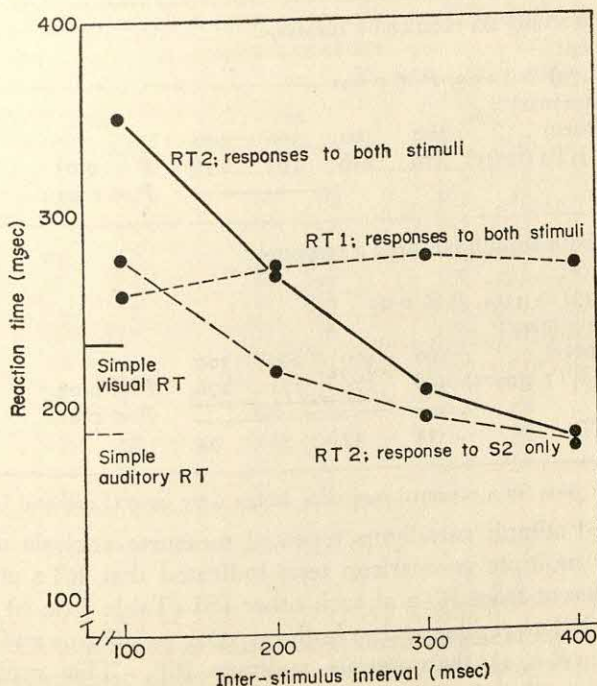
FIGURE 1. Reaction time as a function of inter-stimulus interval with and without a response to the first stimulus of the pair.

The only RTs not included in the computations were obvious or declared anticipations, or very long RTs where subjects clearly inhibited their response. These amounted to a mean of 2·75 rejected RTs in the two response condition and 1·0 in the one response condition, with an expected tendency for number of anticipations to increase with ISI.

In both paired stimuli conditions there was a negatively accelerated curve of improvement of RT to S2 with increasing ISI (Fig. 1), the slope of the curve being much shallower in the one response condition. The finding of a slowed RT to S2 in the one response condition parallels that of Davis (1958) in a similar experiment with highly practised subjects. He hypothesized that this was due to the processing of S1 occupying the single channel mechanism, even when no response was demanded.

TABLE I

Analysis of variance between mean response times to S1 and S2 at different ISIs in the two response and one response paired stimuli conditions

(a) RT2 when both stimuli required a response.
Analysis of variance:
  $F = 124·7$, $F_{(1,33)} \geqslant 12·05$, $P \leqslant 0·001$
Newman–Keuls summary:

ISI (msec)	100	200	300	400	
Mean RT2 (msec)	349	266	205	177	$P < 0·01$

(b) RT to S2 when only S2 required a response.
Analysis of variance:
  $F = 77·19$, $F_{(1,33)} \geqslant 12·05$, $P \leqslant 0·001$
Newman–Keuls summary:

ISI (msec)	100	200	300	400	
Mean RT2 (msec)	276	218	193	173	$P < 0·05$
					$P < 0·01$

(c) RT1s when both stimuli required a response.
Analysis of variance:
  $F = 4·92$, $F_{(1,33)} \geqslant 4·14$, $P \leqslant 0·05$
Newman–Keuls summary:

ISI (msec)	100	400	200	300	
Mean RT1 (msec)	258	270	271	276	$P < 0·05$
					$P \leqslant 0·01$
S.D. RT1	32	37	33	34	

Conservative F tests for a repeated measures design were used (Lana and Lubin, 1963).

In both paired stimuli paradigms repeated measures analysis of variance and Newman–Keuls multiple comparison tests indicated that RT2 at each ISI was significantly different from RT2 at each other ISI (Table I (a, b)). Mean RTs at the 400 msec ISI were identical in both response paradigms and 10 msec faster (non-significant: $F < 1$) than simple auditory RT. This replicated Davis's (1957, 1958) finding that there was no refractoriness at this ISI. However, interindividual comparisons (see Tables II and III) showed that in the two response

TABLE II

Subject mean RT2s and their S.D.s at the different ISIs in the two response paired stimuli condition

ISI Subject	100		200		300		400	
	RT2	S.D.	RT2	S.D.	RT2	S.D.	RT2	S.D.
MM	387	94	268	72	208	81	137	53
PM	302	27	227	32	189	80	170	54
RB	337	52	266	54	218	90	198	71
JHa	307	26	241	53	179	96	207	65
JW	272	71	174	66	126	52	132	104
EC	327	108	229	73	158	82	159	85
SM	323	39	247	115	156	64	167	37
CB	319	48	218	82	159	91	103	60
SL	355	66	250	55	162	75	111	94
DR	382	42	344	81	298	62	247	120
JHi	539	43	459	43	411	52	354	93
AB	344	69	279	51	205	78	147	79
Mean	349	57	266	64	205	75	177	76
S.D.	66		69		76		67	

TABLE III

Subject mean RTs to S2 and their S.D.s at the different ISIs in the one response paired stimuli condition

ISI Subject	100		200		300		400	
	RT	S.D.	RT	S.D.	RT	S.D.	RT	S.D.
MM	362	97	280	88	265	85	200	51
PM	223	90	163	57	143	33	135	18
RB	243	72	175	64	185	56	151	27
JHa	235	64	189	99	136	38	150	29
JW	241	73	151	48	145	45	124	13
EC	287	115	227	66	189	39	184	50
SM	208	47	182	61	153	39	154	16
CB	245	58	200	49	180	23	190	35
SL	302	76	261	101	214	98	206	60
DR	237	58	211	66	180	38	168	23
JHi	408	77	368	34	342	51	308	57
AB	325	84	215	59	184	51	162	37
Mean	276	75	218	66	193	49	177	34
S.D.	61		60		57		45	

condition subject mean RT2s at the 400 msec ISI ranged from 103 msec faster to 53 msec slower (130 msec slower in the case of one subject with atypically long RTs generally) than simple auditory RT. It was also the case that, comparing the two response and one response condition at the 400 msec ISI, subject mean RT2s

ranged from 105 msec faster in the two response condition than the one response condition to 54 msec faster in the one response condition than the two response condition.

An analysis of variance and multiple comparison test on RT1s across the range of ISIs yielded the interesting finding that RT1 was significantly faster (Table I (c)) at the 100 msec ISI than at the other three ISIs, apparently an instance of inter-modality facilitation similar to that demonstrated by Morrell (1967).

The overall mean RT1 of 269 msec was slower than the mean simple visual RT of 233 msec. This is at variance with the results of Davis (1956, 1957) but similar to those of Smith (1967b) and may have been due to the necessity of shifting focal attention between sensory modalities concurrently with responding to S1.

Evoked Potentials to Single Stimuli

Evoked potentials to single stimuli in the control conditions morphologically resembled those recorded by other researchers (e.g. see Bergamini and Bergamasco, 1967; Wilkinson, 1967). Typical visual and auditory EPs are shown in Figure 2.
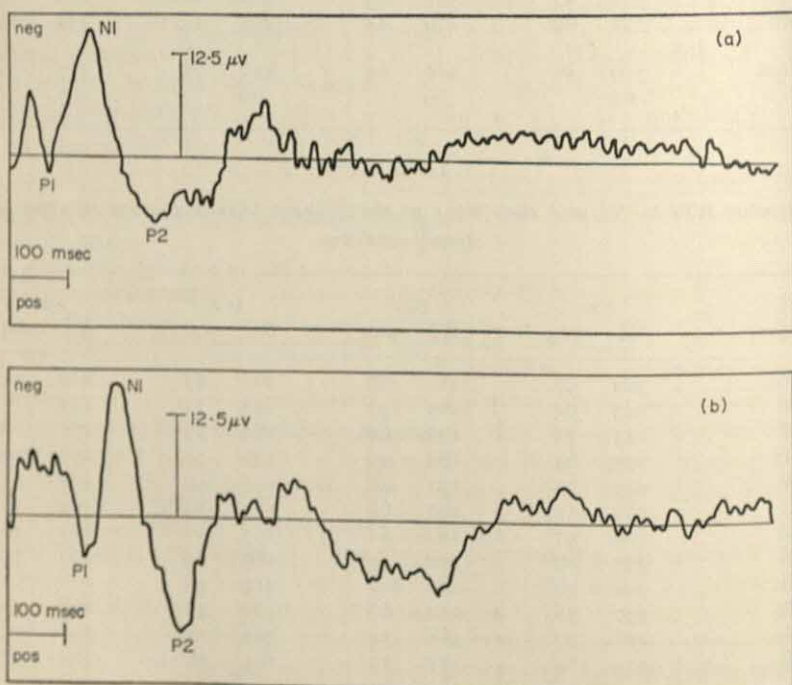
FIGURE 2. The morphology of auditory and visual evoked potentials recorded from the scalp derivation vertex to mastoid.

In the visual EP the prominent initial component of the non-specific response investigated in this experiment was composed of a positive-negative-positive wave (P1–N1–P2 in the nomenclature used by Wilkinson, 1967) with successive maxima and minima having overall mean latencies of 121, 174 and 264 msec. The equivalent peaks in the auditory EP (jointly referred to as the vertex or "V

potential" by Davis *et al.*, 1966) had mean latencies of 64, 123 and 208 msec. Amplitude measurements were made between the peaks P_1-N_1 and N_1-P_2 in the convention used by other experimenters (e.g. Davis *et al.*, 1966). Amplitude measurements on single stimulus EPs are shown in Table IV.

TABLE IV

Amplitude measurements (μV) on single stimulus EPs with and without a response required

		Single stimulus/no response		Single stimulus/simple RT	
		P_1-N_1	N_1-P_2	P_1-N_1	N_1-P_2
Visual	Mean	8·5	17·4	13·9	20·0
	S.D.	4·7	6·3	6·3	6·6
Auditory	Mean	9·8	17·8	15·5	23·3
	S.D.	5·0	9·0	3·0	7·8

In Table IV it can be seen that, for both stimulus modalities the mean amplitudes of both P_1-N_1 and N_1-P_2 were augmented when the stimuli demanded a response, compared with the instance where they were passively observed by the subjects. The P_1-N_1 differences were statistically significant in both modalities (auditory: Sandler's $A^* = 0.153$; visual $A = 0.152$; $A(N = 12) \leqslant 0.178$, $P \leqslant 0.01$) but the N_1-P_2 differences were not significant (auditory: $A = 0.417$; visual: $A = 0.407$; $A(N = 12) \leqslant 0.273$, $P \leqslant 0.05$). This finding concurs specifically with that of Wilkinson and Morlock (1967) and is consistent with other findings associating increased EP secondary response amplitudes with enhanced attention and performance.

Latencies of EP Components across Different ISIs

Identification of the EPs to S_1 and S_2 in serial presentation was accomplished by superimposing tracings of EPs to single stimuli, for the same subject, and judging on the basis of an overall morphological resemblance (see Fig. 3). There was clearly a subjective element in the judgement. As EP duration is up to 400 msec, the ISI range of 100–400 msec implied that the EP to S_1 would overlap that to S_2. As there is some evidence that algebraic summation occurs in these circumstances (Donchin, Wicke and Lindsley, 1963; Walter, 1964) identification of EP2s, with ISIs in the range 100–400 msec, is difficult and comparison on an amplitude basis complicated.

No evidence could be found of increases in latency of non-specific EP components, consistent with delays in RT2. Attempts to identify EP2 indicated that like components had similar latencies to their counterparts in single stimulus EPs, as in non-behavioural experiments on EP recovery functions (Davis *et al.*, 1966). EP2 was almost impossible to identify at the 100 msec ISI (1/12 identifications in the two response condition and 2/12 in the one response condition). EP2 was identified more frequently overall in the one response condition than the two response condition ($\chi^2 = 3.86$; $\chi^2(df = 1) > 3.84$, $P < 0.01$).

---

*Use of Sandler's A tables enables a matched pairs t-test to be performed with a simple computational procedure.
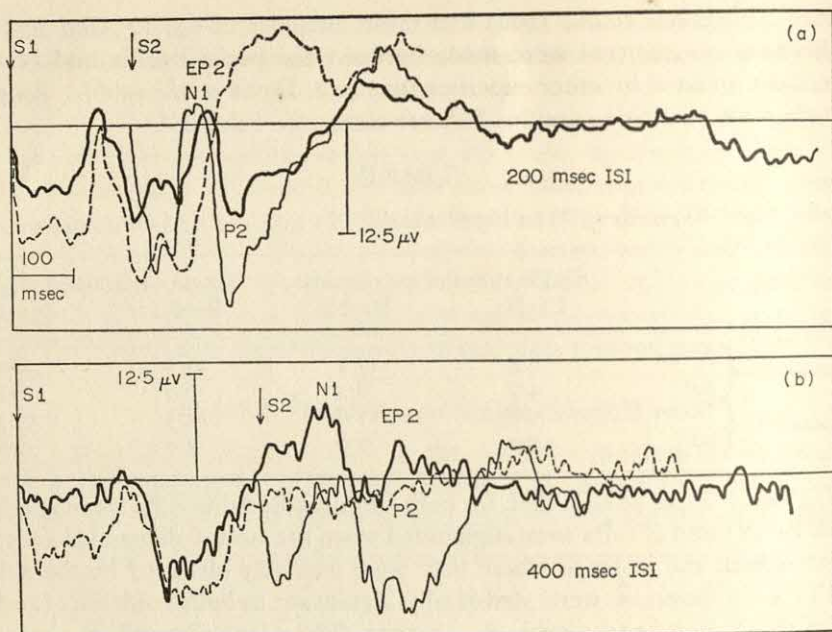
FIGURE 3.   Identification of EP2 in compound EPs produced by visual–auditory stimulus pairs.

In the two response condition 11/12 EP2s were identified at the 200 msec ISI, 8/12 at the 300 msec and 7/12 at the 400 msec ISI.   In the one response condition identification of EP2 was possible for every subject at ISIs of 200 msec or more. The EP2 N1 latencies in the one response condition were calculated and comparison of means across ISIs (Table V (a)) showed no significant difference ($F < 1$). Similar comparisons were made of P1 latencies (Table V(b)) and again no differences were apparent ($F < 1$).   Mean latencies of N1 in the two stimulus condition were all about 10 msec shorter than the latencies of the same component in the EP to single stimuli.   This difference was not significant when tested with a conservative F test (see Table V (c)).

EP Amplitude Refractoriness and Psychological Refractoriness

As no changes in latency specifically associated with psychological refractoriness could be demonstrated, there was further investigation into whether EP amplitude refractoriness correlated with psychological refractoriness.   It was not possible to investigate this adequately on the basis of comparing EP2 amplitudes at different ISIs, as it appears likely that there is algebraic summation or some other complex interaction between EP1 and EP2.   Facilities were not available for the subtraction of an EP to a single stimulus synchronized with EP1 which, on the challengeable assumption of algebraic summation, would have left an uncontaminated EP2.

In fact there was a decrement in the mean uncorrected P1–N1 amplitude between the 200 msec ISI and the 300 and 400 msec ISIs (Table VI).   This can tentatively be ascribed to coincident negative going components of EP1 summing with P1–N1 of EP2 at the 200 msec ISI.   However, one might have expected compensating P1–N1 augmentation at the longer ISIs due to summation with

TABLE V

Latencies of non-specific EP components in different stimulus-response paradigms and at different ISIs within the paired stimulus paradigm (each mean value is the mean of 12 subject means)

(a) Mean N_1 latencies in EP2 at different ISIs, in paired stimulus, one response condition, and mean N_1 latency in single stimulus auditory EPs.
Analysis of variance: $F < 1$

	Simple auditory RT	200	300	400
ISI (msec)		200	300	400
Mean latency (msec)	123	112	112	109
s.d. latency (msec)	9	16	21	19

(b) Mean P_1 latencies in EP2 at different ISIs, in paired stimulus, one response condition, and mean P_1 latency in single stimulus auditory EPs.
Analysis of variance: $F < 1$

	Simple auditory RT	200	300	400
ISI (msec)		200	300	400
Mean latency (msec)	64	62	64	66
s.d. latency (msec)	12	11	14	16

(c) Latin square analysis of variance across overall mean auditory EP N_1 latencies in different stimulus response paradigms.
Analysis of variance:

Conditions	$F = 2.57, F(1, \infty) \geqslant 3.84, P \leqslant 0.05$			
Order	$F < 1$			
Interaction	$F < 1$			
	2 stimulus 1 response	1 stimulus no response	Simple RT	2 stimulus 1 response
Mean latency	111	120	123	124

later negative going components, greater recovery from refractoriness, and because of association with speeded RT. A similar comparison of EP2 N_1–P_2 amplitudes across ISIs yielded no significant differences (Table VI).

Comparisons Between Stimulus–Response Paradigms

A number of other comparisons give a more reliable indication of the fate of EPs to S2 at brief ISIs. RT2s at the 400 msec ISI in the paired stimuli condition were slightly (not significantly) faster than RTs to single auditory stimuli (Fig. 1), indicating that any psychological refractoriness has dispersed at this ISI. If there is associated dispersion of physiological amplitude refractoriness then there should be equality in amplitude of EP components comparing the conditions referred to. This is emphatically not the case as EP2 P_1–N_1 and N_1–P_2 amplitudes at the 400 msec ISI in both paired stimulus conditions are a small fraction of the amplitude of the same deflections in the simple auditory RT EP: highly significant differences ($P < 0.001$ in every case; see Table VII). This amplitude comparison should be

TABLE VI

Analyses of variance of amplitudes of EP2 components across different ISIs in the two stimulus one response condition

(a) P1–N1 amplitudes at different ISIs.
Analysis of variance:
$F = 5\cdot48, F(1,20) \geqslant 4\cdot35, P \leqslant 0\cdot05$
Newman-Keuls summary:

ISI	200	300	400	
Amplitude (μV)	10·6	5·7	7·3	$P < 0\cdot05$
s.d. amplitude	5·6	2·3	3·6	

(b) N1–P2 amplitudes at different ISIs.
Analysis of variance:
$F = 2\cdot54, F(1,22) \geqslant 4\cdot30, P \leqslant 0\cdot05$
Summary data:

ISI	200	300	400
Amplitude (μV)	11·4	8·9	11·7
s.d. amplitude	5·7	4·2	3·6

Conservative F tests for a repeated measures design were used (Lana and Lubin, 1963).

reliable as the non-specific components of EP2 are uncontaminated by EP1 beyond 400 msec. Furthermore, it was found in another study by Boddy (1970) that there was still profound amplitude attenuation of EP2 at an ISI of 1·25 sec. In this case an RT response to S2 was facilitated by the preparatory adjustments initiated by S1.

TABLE VII

P1–N1 and N1–P2 amplitudes in EP2s at the 400 msec ISI in the paired stimuli conditions and EPs from simple auditory RT trials

		(1) 1 stimulus 1 response	(2) 2 stimulus 1 response	(3) 2 stimulus 2 responses
P1–N1	Mean (μV)	15·5	7·3	4·6
	s.d.	3·8	5·2	5·8
N1–P2	Mean (μV)	23·3	11·7	3·7
	s.d.	7·8	3·6	4·5

Significance of differences using Sandler's A test:

P1–N1	1 and 2	$A = 0\cdot107$	
	1 and 3	$A = 0\cdot100$	$A(N = 12) \leqslant 0\cdot130,$
N1–P2	1 and 2	$A = 0\cdot124$	$P \leqslant 0\cdot001$
	1 and 3	$A = 0\cdot097$	

The findings reported combine to suggest that even where evoking stimuli demand an urgent behavioural response, EP recovery may take little shorter time than the upper limit of 10 sec found in experiments by Allison (1962) and Davis et al. (1966) where there was passive reception of stimuli.

A point of criticism should be anticipated. In the experiment reported here single stimulus EPs were derived from sequences of 16 stimuli, whilst paired stimulus EPs were derived from approximately 16 stimuli randomly distributed in a sequence of 64. It could be argued that the attenuation observed in the latter case was due to progressive habituation or refractoriness over the longer overall sequence of stimuli. Reference to the careful work of Ritter *et al.* (1968) on the habituation of EPs indicates that these factors could only account for a very small proportion of the attenuation observed, if any, at the inter-pair intervals used. This apart it would still be inconsistent with previous findings for the behavioural performance to be maintained whilst the EP habituated.

Thus it appears that psychological amplitude refractoriness persists far beyond the time when recovery from physiological refractoriness is complete and even into the range where behavioural performance is enhanced. Furthermore, this instance where physiological amplitude refractoriness is induced in a behavioural situation defines one circumstance where the classical relationship associating augmented non-specific EP components with enhanced behavioural performance breaks down.

In two comparisons the usual relationship between EP amplitude and performance is actually inverted. Five subjects responded substantially faster to S2 at the 400 msec ISI in the two response condition than in the simple auditory RT condition (mean difference 76 msec, S.D. 20, range 37–103), yet the amplitudes of P_1–N_1 and N_1–P_2 in the EPs associated with the faster RTs are either missing (one case) or for all the other subjects highly attenuated (P_1–N_1 by a mean of 55% and N_1–P_2 by a mean of 77%) relative to the amplitudes of these deflections in the single stimulus auditory EP associated with the slower RT response.

In another comparison five subjects responded faster to S2 at the 400 msec ISI, when a response was required to both stimuli, than when a response was required to S2 only (mean difference 58 msec, S.D. 33, range 15–105). However, in every individual comparison except one (the subject with the smallest RT difference) the amplitudes of both P_1–N_1 and N_1–P_2 were again greater in the EP associated with the slower RT (mean differences 3·1 and 6·7 μV, S.D.s 3·0 and 5·1, ranges 0·1–6·8 and −2·3–12·5 μV).

EP Refractoriness Induced by Behavioural Response

The comparison cited above leads to another important finding comparing the EP2s in the one response and two response paired stimulus conditions at the 400 msec ISI. The mean P_1–N_1 and N_1–P_2 amplitudes in the one response condition (7·3 μV, S.D. 5·1 and 11·6 μV, S.D. 3·6) were greater than those in the two response condition (mean P_1–N_1 4·6 μV, S.D. 5·8 and mean N_1–P_2 3·7 μV S.D. 4·5). The difference was not significant in the case of P_1–N_1 (Sandler's $A = 0·327$; $A(N = 12) \leqslant 0·273$, $P \leqslant p·05$) but was highly significant in the case of N_1–P_2 ($A = 0·114$; $A(N = 12) \leqslant 0·130$, $P \leqslant 0·001$).

The inference to be drawn from this finding is that the occurrence of a behavioural response to S1 is a source of amplitude refractoriness in S2 in addition to that induced by S1 alone. Furthermore, the amplitude attenuation induced by this source of refractoriness has no counterpart in slowed RT performance as this was exactly the same in the two conditions compared.

Correlative Studies

If the undulations of the EP reflect the time course of information processing in underlying neural tissue then one might expect inter-individual correlations between speed of performance and the latencies and periodicities of the relevant components of the EP. The Pearson product moment correlation coefficient between N_1 latency and simple RT was 0·06 for auditory stimuli (based on a group of 24 subjects as data from a further 12 subjects used in another EP experiment was available) and 0·04 (12 subjects) for visual stimuli, clearly non-significant in each case. The correlation between the P_1–P_2 period and RT was 0·22 (24 subjects: $r(df = 25) \geqslant 0·32$, $P \leqslant 0·10$) for auditory stimuli and 0·01 for visual stimuli. There was a similar lack of correlation between the P_1–N_1 period for both stimulus modalities ($r = 0·00$ in each case).

The primary concern in this experiment was with delays in the second RT when it closely followed the first. Evidence was sought that the delays in RT_2 were related to the time course of EP_1, whose appearance supposedly reflects information processing blocking a central channel. RT_2 delay was calculated by subtracting simple auditory RT from RT_2 at the 100 msec ISI, in the two response condition, for each subject. 100 msec ISI data was used as RT_2 delay was most marked in this instance. As the evidence pointed to stability in EP component latencies, latency data from EPs to single visual stimuli were correlated with RT_2 delays. These EPs were more clearly defined because there was no superposition of a second EP, and because the data was more complete. The highest correlation was that between RT_2 delay and P_1 latency ($r = 0·39$; $r(df = 9) \geqslant 0·52$, $P \leqslant 0·10$). Correlations for N_1 latency and P_1–P_2 period were 0·02 and 0·21 respectively. Two subjects were excluded from these computations because of poor EP data. Thus no evidence could be found that EP_1 reflected the time course of information processing which was delaying further input to a single channel central mechanism.

The only significant correlations occurred in investigation of the seemingly tenuous possibility of a systematic relationship between RT_2 at the 100 msec ISI, in the two response condition, and the time course of auditory EPs. Again single stimulus EP data were used. There were inter-individual correlations of 0·59 between P_2 latency and RT_2 delay and 0·62 between P_1–P_2 period and RT_2 delay (confidence level: $r(df = 9) \geqslant 0·60$, $P \leqslant 0·05$). The highest correlation was 0·65 between RT_2 and P_1–P_2 period. However, there were no significant correlations in the condition where only S_2 demanded a response, the correlations being 0·35 between RT_2 delay and P_1–P_2 period, 0·06 between RT_2 delay and P_1 latency and 0·18 between RT_2 delay and N_1 latency. Interpretation of the significant correlations is tackled in the discussion section.

Discussion

There was no evidence of an increase in latency of non-specific EP components associated with "psychological refractoriness", exhibited in the delayed RT

response to the second of two close stimuli. Thus the temporal course of the EP to the second stimulus did not appear to reflect the operation of a hypothesized single channel information processor delaying the admission of signals arriving before the termination of an operation in progress. The implications of this finding are either that the single channel hypothesis is wrong or that EPs do not reflect the time course of the information processing which occurs.

To deny the correctness of the single channel hypothesis would challenge an impressive body of supporting evidence. However, in a PRP experiment by Boddy (1970) using 20 subjects, it was found that subjects responding relatively slowly to S1 responded relatively fast to S2 and vice versa. This finding could not be attributed to occasional grouping of responses, supposed to be made separatley to each of the stimuli. The negative correlation suggested that individuals varied in the emphasis which they placed on responding rapidly to the first stimulus relative to the second and this seemed to presuppose the possibility of operating in a multi-channel mode. In a single channel system a relatively fast response to S2 would be inconsistent with the delay imposed by a relatively slow response to S1. The appearance of single channel operation may have been a function of the constraints placed upon the subjects in the classical experiments. The possibility of multi-channel operation could be systematically investigated by manipulating the relative priorities attached to responding to each of the stimuli in a PRP experiment.

On a preliminary examination the evoked potential data from the reported experiment support the hypothesis that there is parallel processing of the two closely succeeding stimuli as EP2, though unchanged in latency, is apparently amplitude attenuated. This would be consistent with the idea that RT2 delays originate from the "sharing" of an overall processing capacity which reduces the efficiency of stimulus evaluation and response organization. However, this argument collapses in the light of the further finding that the amplitude refractoriness persists after recovery from psychological refractoriness is complete.

A further interesting finding was that there was greater amplitude refractoriness in EP2 when S1 commanded a motor response than when it was passively received, indicating that the motor response was itself a source of refractoriness. Gilden, Vaughan and Costa (1966) have shown that spontaneous voluntary muscle contractions are accompanied by a scalp potential which they named the "motor potential", and which Wilkinson (1967) has suggested is compound in the late stages (peaks at 330 and 400 msec) of EPs to RT stimuli. It may therefore be that the sections of the EP attributable to the stimulus and to the motor response respectively are additive sources of physiological refractoriness.

The concomitant of the finding of a dissociation between psychological and physiological refractoriness was that the experimental design demonstrating it produced a breakdown of the usual positive correlation between EP amplitude and indices of performance, and sometimes even a reversal. This finding indicates caution in the interpretation of the classical relationship in terms of attentional mechanisms.

Only ad hoc formulations rescue the hypothesis that the non-specific EP reflects information processing from discredit. One is that S1 initiates the central and

peripheral organization of a response to S2, for which S2 is merely a "trigger" requiring minimal evaluation. The minimal processing required for S2 may reduce its vulnerability to physiological refractoriness. In Woodworth and Schlosberg (1954) it is suggested that rising muscular tension in RT foreperiods reflects an incipient response movement. The hypothesis could only be satisfactorily investigated by subjectively dissociating a prior stimulus and an RT signal, which would be difficult to achieve.

There is a further ad hoc formulation which partially preserves the integrity of the initial hypotheses concerning both the psychological and the physiological phenomena. Goff (1969) has argued that non-specific EP components and conscious awareness are intimately related. It may be that EPs are primarily a manifestation of information processing associated with conscious identification, evaluation and response organization on the occurrence of novel and/or complex stimuli. In other words EPs may reflect the operation of a standard evaluative and integrative "subroutine", which is essentially superfluous during the performance of simple and highly practised skills like RT, probably monitored at a lower level in the nervous system. Work associating EPs with more complex evaluations, like stimulus form (John et al., 1967) and stimulus uncertainty (Sutton, Tueting, Zubin and John, 1967), have been the basis for suggesting that EPs reflect information processing.

In the experiment reported there was an inter-individual correlation between the period of the major P_1–N_1–P_2 wave and both RT2 delay ($r = 0.62$) and RT2 ($r = 0.65$) at the 100 msec ISI, when responses were required to both stimuli. This represented the instance when the subjects' task was most demanding in terms of rapid stimulus evaluation and response organization, and thus may have been sufficiently complex to involve the "conscious evaluation subroutine" which has been postulated, so that the associated EP reflected the time course of information processing whose duration determined response latency. Use of pairs of stimuli calling for discriminations and decisions according to complex criteria relating to the stimulus form could be used to test the hypothesis presented. Its validity would be indicated by a close relation between performance recovery and EP recovery with increasing ISI.

In conclusion it has been shown that there is a dissociation between psychological refractoriness and "physiological refractoriness" expressed in terms of either latency or amplitude. Furthermore, in the stimulus response paradigm used, the usual relationship of augmented EP and speeded behavioural performance broke down, indicating interpretive caution. However, it is suggested that a search for a relationship between EPs and more complex behaviour than RT to undifferentiated stimuli might be more rewarding.

References

ALLISON, T. (1962). Recovery functions of somatosensory evoked responses in man. Electroencephalography and Clinical Neurophysiology, **14**, 331–43.

BERGAMINI, L. and BERGAMASCO, B. (1967). Cortical Evoked Potentials in Man. Springfield, Ill.: Charles C. Thomas.

BERTELSON, P. (1966). Central intermittency twenty years later. *Quarterly Journal of Experimental Psychology*, **18**, 153–63.

BICKFORD, R. G., GALBRAITH, R. F. and JACOBSON, J. L. (1963). The nature of averaged evoked potentials recorded from the human scalp. *Electroencephalography and Clinical Neurophysiology*, **15**, 720.

BICKFORD, R. G., JACOBSON, J. L. and CODY, T. R. (1964). Nature of averaged evoked potentials to sound and other stimuli in man. *Annals of the New York Academy of Sciences*, **112**, 204–18.

BODDY, J. (1970). The behavioural significance of some EEG phenomena. Unpublished Ph.D. thesis, University of Manchester, England.

BRAZIER, M. A. B. (1964). Evoked responses recorded from the depths of the human brain. *Annals of the New York Academy of Sciences*, **112**, 33–59.

CIGANEK, L. (1964). Excitability cycle of the visual cortex in man. *Annals of the New York Academy of Sciences*, **112**, 241–53.

CRAIK, K. J. W. (1947). Theory of the human operator in control systems. I. The operator as an engineering system. *British Journal of Psychology*, **38**, 56–61.

CRAIK, K. J. W. (1948). Theory of the human operator in control systems. II. Man as an element in a control system. *British Journal of Psychology*, **38**, 142–8.

DAVIS, H., ENGEBRETSON, H., LOWELL, E. L., MAST, T., SATTERFIELD, J. and YOSHIE, N. (1964). Evoked responses to clicks recorded from the human scalp. *Annals of the New York Academy of Sciences*, **112**, 224–5.

DAVIS, H., MAST, T., YOSHIE, N. and ZERLIN, S. (1966). The slow response of the human cortex to auditory stimuli: Recovery process. *Electroencephalography and Clinical Neurophysiology*, **21**, 105–13.

DAVIS, R. (1956). The limits of the "psychological refractory period". *Quarterly Journal of Experimental Psychology*, **8**, 24–38.

DAVIS, R. (1957). The human operator as a single channel information system. *Quarterly Journal of Experimental Psychology*, **9**, 119–29.

DAVIS, R. (1958). The role of "attention" in the psychological refractory period. *Quarterly Journal of Experimental Psychology*, **10**, 211–20.

DAVIS, R. (1962). Choice reaction times and the theory of intermittency in human performance. *Quarterly Journal of Experimental Psychology*, **14**, 157–66.

DAVIS, R. (1965). Expectancy and intermittency. *Quarterly Journal of Experimental Psychology*, **17**, 75–8.

DOMINO, E. F. and CORSSEN, G. (1964). Visually evoked response in anaesthetized man with and without induced muscle paralysis. *Annals of the New York Academy of Sciences*, **112**, 226–37.

DOMINO, E. F., MATSUOKA, S., WALTZ, J. and COOPER, I. (1964). Simultaneous recording of scalp and epidural somatosensory evoked responses in man. *Science, New York*, **145**, 1199–1200.

DONCHIN, E., WICKE, J. D. and LINDSLEY, D. B. (1963). Cortical evoked potentials and perception of paired flashes. *Science*, **141**, 1285–6.

DONCHIN, E. and COHEN, L. (1967). Averaged evoked potentials and intra-modality selective attention. *Electroencephalography and Clinical Neurophysiology*, **22**, 537–46.

GILDEN, L., VAUGHAN, H. G. and COSTA, L. D. (1966). Summated human EEG potentials with voluntary movement. *Electroencephalography and Clinical Neurophysiology*, **20**, 433–8.

GOFF, W. R. (1969). Evoked potential correlates of perceptual organization in man. In EVANS, C. R. and MULHOLLAND, T. B. (Eds), *Attention in Nureophysiology*, pp. 169–93. London: Butterworths.

GOLDSTEIN, M. H. (1961). Averaging techniques applied to evoked responses. *Electroencephalography and Clinical Neurophysiology*, Suppl. **20** (Computer techniques in EEG Analysis), 59–63.

HAIDER, M. (1967). Vigilance, attention, expectation and cortical evoked potentials. *Acta Psychologica*, **27**, 246–52.

JOHN, E. R. (1967). *Mechanisms of Memory.* New York & London: Academic Press.

JOHN, E. R., HERRINGTON, R. N. and SUTTON, S. (1967). Effects of visual form on the evoked response. *Science,* **155,** 1439–42.

KATZMAN, R. (1964). The validity of the visual evoked response in man. *Annals of the New York Academy of Sciences,* **112,** 238–40.

LEVONIAN, E. (1966). Evoked potential in relation to subsequent alpha frequency. *Science,* **152,** 1280–2.

LANA, R. E. and LUBIN, A. (1963). The effect of correlation on the repeated measures design. *Educational and Psychological Measurement,* **23,** 729–39.

MORRELL, L. (1967). Intersensory facilitation of reaction time. *Psychonomic Science,* 8, 77–8.

ORNITZ, E. N., RITVO, E. R., CARR, E. M., LA FRANCHI, S. and WALTER, R. D. (1967). The effect of sleep onset on the auditory evoked response. *Electroencephalography and Clinical Neurophysiology,* **23,** 335–41.

RITTER, W., VAUGHAN, H. G. and COSTA, L. D. (1968). Orienting and habituation to auditory stimuli: A study of short term changes in averaged evoked responses. *Electroencephalography and Clinical Neurophysiology,* **25,** 550–6.

SMITH, M. C. (1967a). Theories of the psychological refractory period. *Psychological Bulletin,* **67,** 202–13.

SMITH, M. C. (1967b). Reaction time to a second stimulus as a function of intensity of the first stimulus. *Quarterly Journal of Experimental Psychology,* **19,** 125–32.

SUTTON, S., TUETING, P., ZUBIN, J. and JOHN, E. R. (1967). Information delivery and sensory evoked potential. *Science,* **155,** 1436–9.

TELFORD, C. W. (1931). The refractory phase of voluntary and associative responses. *Journal of Experimental Psychology,* **14,** 1–36.

VAUGHAN, H. G. (1966). The perceptual and physiologic significance of visual evoked responses recorded from the scalp in man. *Vision Research,* 6, Suppl. 1, 203–23.

WALTER, W. G. (1964). The convergence and interaction of visual auditory and tactile responses in the non-specific human cortex. *Annals of the New York Academy of Sciences,* **112,** 320–61.

WEITZMAN, E. D. and KREMEN, H. (1965). Auditory evoked responses during different stages of sleep in man. *Electroencephalography and Clinical Neurophysiology,* **18,** 65–70.

WILKINSON, R. T. (1967). Evoked responses and reaction time. *Acta Psychologica,* **27,** 235–45.

WILKINSON, R. T. and MORLOCK, H. C. (1967). Auditory evoked response and reaction time. *Electroencephalography and Clinical Neurophysiology,* **23,** 50–6.

WOODWORTH, R. S. and SCHLOSBERG, H. (1954). *Experimental Psychology.* New York: Holt.

INTERPRETATION AND MATCHING BIAS IN A REASONING TASK

J. ST. B. T. EVANS

Psychology Section, City of London Polytechnic

Subjects were required to construct verifying and falsifying cases of conditional rules in which the presence and absence of negative components was varied. Their responses gave some indication of consistent interpretation of the rules, generally conforming to Wason's (1966) idea of a "defective" truth table. Much of their behaviour, however, seemed to be determined by a task variable in the form of a tendency to construct instances which matched, rather than altered, the values named in the rules.

Introduction

Most studies of propositional reasoning have required subjects to make inferences with rules of the form, *If p then q*. This investigation is concerned with discovering the way in which subjects understand and interpret such rules, and deciding whether or not their interpretation is altered by the presence of negative components. To this end an attempt is made to elicit the psychological "truth table" which subjects possess for each rule.

The truth table for conditional rules, within the propositional calculus of logic, is stated in terms of the truth and falsity of each component (antecedent and consequent). For simplicity consider the affirmative example, "If it is a football then it is round". Logically this rule could only be falsified by finding a football which is not round. This is known as the TF (true-false) case, since the antecedent is true and the consequent false. The rule is obviously verified by a round football (the TT case) but what of objects which are not footballs whether round (FT) or not (FF)? Within the calculus they are considered to verify the rule, but Wason (1966) suggested that subjects consider such cases to be irrelevant, and are consequently in possession of a "defective" truth table for a conditional rule.

Johnson-Laird and Tagart (1969) conducted an experiment in which subjects were required to classify instances according to whether they verified (T), falsified (F) or were irrelevant (?) to a given rule. With rules of the form, *If p then q*, they found strong support for Wason's hypothesis of a "defective" truth table: 19 of the 24 subjects classified the cases TT, TF, FT and FF as T, F, ?, and ? respectively. It might be argued, however, that by giving the category "irrelevant", the experimenters were suggesting to the subjects that they should make use of a third truth value. The authors refuted this suggestion on the grounds that this category was infrequently used when the rule was presented in different terms (e.g. a disjunctive). In the present study, however, subjects were asked to construct as many verifying and falsifying cases of a given rule as they could think of, so that any unused case might, with some justification, be inferred by the experimenter as "irrelevant".

A second problem in the Johnson-Laird and Tagart study is common also to a number of studies of deductive reasoning. Although they used negative components in some of the rules, all forms expressed the same logical relationship, *p implies q*, so that the truth and falsity of components was inevitably confounded with affirmation and negation. The TT, TF, FT and FF cases of such rules are produced by the combinations, pq, pq̄, p̄q and p̄q̄, where p̄ = *not p* and q̄ = *not q*. The problem can be overcome, however, if negative components are systemically varied with affirmatives in "If . . . then . . ." sentences (Table 1). Now each combination of values corresponds to a different truth case on each rule. Hence, for example, pq̄ constitutes the case TF for the rule, *If p then q*, TT for *If p then not q*, FF for *If not p then q* and FT for *If not p then not q*. Thus overall the effect of instances matching (affirming) or mismatching (negating) values named in the rules should cancel out.

TABLE I

The combinations of affirmed and negated values constituting the different truth table cases of the conditional rules used in the experiment

| Rule | Truth table cases | | | |
	TT	TF	FT	FF
(1) If p then q	pq	pq̄	p̄q	p̄q̄
(2) If p then not q	pq̄	pq	p̄q̄	p̄q
(3) If not p then q	p̄q	p̄q̄	pq	pq̄
(4) If not p then not q	p̄q̄	p̄q	pq̄	pq

It is possible, of course, that some difference of interpretation between the rules may be reflected in the truth tables elicited. In an earlier study (Evans, 1972) it was found that subjects made significantly fewer valid *Modus Tollens* inferences with rules which had negated antecedents. Given the premises, (1) *If p then q* and (2) *not q*, for example, most subjects inferred, by *Modus Tollens*, the conclusion, *not p*, but given (1) *If not p then q* and (2) *not q*, few were able by the same inference to conclude *p*. In the latter case most subjects considered the truth of the antecedent, *not p*, to be "indeterminate". Matalon (1962) has shown that the ability to make *Modus Tollens* and the ability to identify TF as a falsifying case are developmentally linked. Since failure to make *Modus Tollens* is also logically inconsistent with the belief that TF falsifies, it was predicted in the present experiment, that fewer TF cases would be constructed to falsify rules which had negative rather than affirmative antecedents.

Method

Design

Each subject was required to construct both verifying and falsifying cases of each of the four rules shown in Table 1. 24 subjects were used, each receiving a different one of the 4! permutations of presentation order. Half of the subjects were assigned randomly to a group who constructed verifying cases before falsifying cases on each rule, the other half always constructing falsifying cases before verifying cases.

Subjects

24 undergraduate students of University College London served as subjects on a paid volunteer basis, and were tested individually.

Task and materials

The subject was presented with an ordered 4 × 4 matrix of stimulus cards depicting figures varying in shape (circle, triangle, cross and square) and in colour (red, yellow, green and blue). Each of the four logical rules was presented to the subject, one at a time, phrased in terms of a relationship between figures in the display. The general form of these rules may be illustrated by the following example of an *If p then not q* rule:

> If there is a red triangle on the left,
> then there is not a green square on the right.

The subject's task was to take two figures from the display and place them side by side in order to make the rule either true or false according to the instruction at the time. If, for example, a subject asked to verify the rule given above, placed a red triangle to the left of (say) a blue circle (i.e. *not* a green square) he would have constructed the TT case. In this manner any pair of cards selected could be classified as constituting one of the four cases TT, TF, FT and FF in relation to a given rule. The lexical materials were varied as much as possible between rules.

Procedure

Each subject was given the following typed instructions to read:

> "You will be presented with a series of rules which always assume that two of the figures before you have been placed side by side. Your task will be to select two of the figures from the array and place them in such a way as to make a given rule true, or in such a way as to make a given rule true, or in such a way as to make a given rule false, according to the instruction. If you have any questions please ask them now and not after you have started on the problems."

The subject was then given a rule and, according to his group, asked to find a way of making it true or making it false. After each solution he was asked to give another instance until he said there were no more. (The subjects were allowed to make generalizations such as "a red triangle on the left with anything other than a green square on the right would make the rule true".) In this way an exhaustive series of verifying and falsifying constructions was obtained for each rule.

Results

(i) An analysis of initial responses

It seems reasonable to suppose that the first verifying and first falsifying construction made on each rule would be the psychologically "strongest", on the grounds that they occurred most immediately to the subject. The frequency with which each truth case was constructed to verify and falsify each rule, on subjects, initial constructions only, is given in Table II.

Table II

The frequency with which subjects gave each truth table case on their initial verification and initial falsification of each rule. (n = 24)

Rule	Verification				Falsification				None
	TT	TF	FT	FF	TT	TF	FT	FF	
1) If p then q	24	0	0	0	0	17	0	6	1
2) If p then not q	24	0	0	0	0	23	0	0	1
3) If not p then q	24	0	0	0	0	7	15	2	0
4) If not p then not q	22	0	0	2	0	13	2	8	1

The most striking feature of the results in Table II is the almost universal construction of the TT case as a verifying instance of all four rules. This is consistent with Wason's (1966) "defective" truth table in which the TT case is the only verifying one. Although the TF case predominates in the initial falsifying constructions, there is, however, considerably more variation.

The hypothesis that fewer TF cases would be constructed to falsify rules with negative antecedents was confirmed ($P = 0.0005$, sign test). There was also a significantly greater number of TF cases constructed to falsify rules with negative consequents ($P = 0.012$, sign test). Taken together, these findings suggest a tendency for subjects to prefer to match rather than alter the values named in the rules, which we shall refer to as "matching bias". Matching values will tend to make affirmative components true and negative components false. Hence, *If p then not q* was the easiest rule to falsify (correctly) since the TF case constitutes the double matching (or affirmation) of components, pq. *If not p then q* was the hardest since the construction of TF requires the double mismatching (or negation) of components, p̄q̄. This interpretation of the TF construction is corroborated by the fact that the most frequent constructions of the FT and FF cases occurred on the *If not p then q* and *If not p then not q* rules respectively, where they constitute in each case the double match, pq.

(ii) *The plotting of psychological truth tables from the total solutions given*

By taking account of all the verifying and falsifying solutions given by each subject on each rule, it was possible to construct three valued truth tables for each rule. If, for example, a subject gave the TT and FT cases to verify and the TF case to falsify a rule, one might infer the case FF to be irrelevant for him to the truth of the rule. The frequencies with which subjects constructed each case as either true or false, or failed to construct it as either ("irrelevant" category), are shown in Table III for each of the four rules.

TABLE III

Three valued psychological truth tables for each rule, constructed by plotting the frequency of true, false and "irrelevant" (non-constructed) classifications of each truth table case
(n = 24)

| Rule | Truth value | Truth table cases | | | |
		TT	TF	FT	FF
(1) If p then q	T	24	0	2	4
	F	0	21	7	8
	?	0	3	15	12
(2) If p then not q	T	24	0	5	10
	F	0	23	1	1
	?	0	1	18	13
(3) If not p then q	T	24	1	2	11
	F	0	15	18	4
	?	0	8	4	9
(4) If not p then not q	T	23	2	4	7
	F	0	18	7	9
	?	1	4	13	8

The effect of "matching bias" on the frequencies recorded in Table III was principally to suppress the construction of some cases which involved mismatching or altering named values, thus increasing their apparent irrelevance. Presumably, this factor would not affect the *direction* of classification (true or false) of those instances which were constructed. A series of sign tests were carried out comparing the frequency of true and false constructions of each case on each rule, ignoring the "irrelevant" responses. On all rules the TT case was significantly more often constructed as true than false ($P < 0.001$), and the TF case was significantly more often constructed as false than true ($P < 0.001$). There were, however, only two significant effects on the false antecedent cases, resulting from a bias to falsity in the construction of FT cases with the rule, *If not p then q* ($P = 0.002$), and a bias to truth in the construction of FF cases of the rule, *If p then not q* ($P = 0.012$).

Discussion

Although the analysis of initial responses clearly supported the hypothesis that fewer correct TF cases would be given to falsify rules with negative antecedents, it cannot be taken as a confirmation of the underlying rationale (consistency with the reasoning observed by Evans, 1972). Along with a number of other trends it is consistent with the idea that subjects prefer to match rather than alter values named in the rules. The detection of a task variable ("matching bias"), correlated with the affirmation and negation of named values, emphasizes the weakness of other studies of deductive reasoning which have focused on rules of the form, *If p then q,*

in which such a factor would be completely confounded with the truth and falsity of the components of the rule. Matching bias may well be a factor in Wason's (1968, 1969) well known "selection task", as Johnson-Laird and Wason (1970) in their information-processing account of Wason's data appear to recognize: ". . . the revised model assumes that the subject without insight will focus on cards mentioned in the rule" (p. 142).

Matching bias does, of course, interact with the subjects, interpretation of the rules in determining their behaviour. Thus from the initial constructions (Table II) it would appear that the TT case is so strongly perceived as a verifying instance as to override completely the effects of matching bias. More influenced by matching, but indicated strongly in the overall analysis (Table III), is the tendency to falsify all rules with the TF case. The construction of FT and FF cases, however, seems to be frequently suppressed when mismatching would result. However, since in general the constructions which were made on these cases were randomly distributed between "true" and "false", we must, like Johnson-Laird and Tagart (1969), infer general support for Wason's (1966) hypothesis of a "defective" truth table.

An inspection of Table III, however, indicates a striking exception in the case of the rule, *If not p then q*. The predominant classifications for the cases TT, TF, FT and FF are T, F, F, and T respectively, which, in the propositional calculus, would be appropriate for a rule of equivalence, *If and only if not p then q*. Can it be that subjects make such an interpretation of this rule which is, in fact, equivalent to the exclusive disjunction, *Either p or q, but not both*? As early as the second century A.D., Galen had observed of conditional rules with negated antecedents that ". . . they are called conditionals by those who pay attention only to the sounds, but a disjunction by those who pay attention to what is meant" (cited by Wason and Johnson-Laird, 1972). If an *exclusive* disjunction is suggested to the subject by the rule, *If not p then q*, perhaps because it is normally used in such a context, then naturally the false antecedent case, *p*, would become relevant since it seems to exclude the possibility of *q*. For example, "If I do not go out then I shall do some work" seems to suggest that going out would prevent me from working.

It is clear from this study that when drawing conclusions about the way in which subjects understand logical rules, one must beware of the influence of task variables. It would seem, for example, that subjects' inability to draw *Modus Tollens* inferences about rules with negative antecedents (Evans, 1972) is, in fact, due to some task difficulty involved in denying a negative statement, since subjects in the present study were generally able to falsify correctly such rules at some stage. (The author follows Legrenzi, 1970, in considering subjects' reasoning to be in error only when it is inconsistent with their evaluation, or in this case construction, of truth table contingencies.) The complexity of interpreting reasoning data is, however, further illustrated in the present study by the fact that while introducing negatives into the rules controlled overall for the effect of matching bias, it appeared, in at least one case, to alter the subjects' interpretation of the conditional sentence.

References

EVANS, J. ST. B. T. (1972). Reasoning with negatives. *British Journal of Psychology*, **63**. In press.

JOHNSON-LAIRD, P. N. and TAGART, J. (1969). How implication is understood. *American Journal of Psychology*, **82**, 367–73.

JOHNSON-LAIRD, P. N. and WASON, P. C. (1970). A theoretical analysis of insight into a reasoning task. *Cognitive Psychology*, **1**, 134–48.

LEGRENZI, P. (1970). Relations between language and reasoning about deductive rules. In D'ARCAIS, G. B. F. and LEVELT, W. J. M. (Eds), *Advances in Psycholinguistics*. Amsterdam, London: North Holland.

MATALON, B. (1962). Etude génétique de l'implication. *Etudes d'épistemologie*, **16**: *Implication, formalisation et logique naturelle.* 69–95.

WASON, P. C. (1966). Reasoning. In Foss, B. M. (Ed.), *New Horizons in Psychology*. Harmondsworth: Penguin.

WASON, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, **20**, 273–81.

WASON, P. C. (1969). Regression in reasoning. *British Journal of Psychology*, **60**, 471–80.

WASON, P. C. and JOHNSON-LAIRD, P. N. (1972). *Psychology of Reasoning: Structure and Content.* London: Batsford. In press.

WITHIN-DAY SERIAL REVERSAL OF A POSITION DISCRIMINATION BY PIGEONS

EUAN M. MACPHAIL

Laboratory of Experimental Psychology, University of Sussex

Naive ($N = 3$) and experienced ($N = 3$) pigeons performed 3 reversals of a simultaneous position discrimination to the same criterion each day. In terms of trials to criterion, or number of correct choices preceding the criterion run, daily first reversals (R1s) were, for both groups, more difficult than either second (R2s) or third reversals (R3s), which did not differ on these measures. The results are interpreted as giving good support to the hypothesis that little or no inhibition is generated in R1s, whereas inhibition does occur in R2s and R3s. An explanation of the relative lack of inhibition in R1s, which relies on concepts drawn from frustration theory, is proposed.

Introduction

Hyperstriatal lesions in birds may disrupt the reversal of a simultaneous discrimination without impairing its original acquisition (Stettner and Schultz, 1967; Macphail, 1971); an implication of this fact is that reversal involves a mechanism that is not essential to acquisition. Theoretical accounts of a similar dissociation induced by mammalian brain lesions have suggested that the function of the damaged mechanism is that of inhibiting responses (e.g. Douglas, 1967). There is at present, however, little behavioural evidence that response inhibition is less involved in the acquisition than in the reversal of simultaneous discriminations. An experiment that gave some indirect support to the proposal that inhibition is not involved in acquisition was reported by Macphail (1970), who showed that pigeons given three serial reversals of a red-green discrimination each day found the second of each day's reversals (R2s) less difficult (in terms of trials to criterion) than either first reversals (R1s) or third reversals (R3s). R1s were regarded as being analogous to conventional acquisition as, due to the occurrence of overnight "forgetting" (Mackintosh, McGonigle, Holgate and Vanderver, 1968), birds showed no preference for either stimulus at the beginning of R1s. Analysis of error patterns showed that the relative ease of R2s was due to the fact that birds performing R2s were less likely to make an error following a run of correct choices than were birds performing R1s or R3s. It was argued that this result was due to the operation of two factors: first, birds performing R1s did not learn to avoid the negative stimulus (owing to the absence of a run of initial errors), and so, after a sequence of correct choices, had to choose between a stimulus that was accumulating approach strength and a stimulus that was neutral rather than (as at a comparable point in R2s or R3s) inhibitory; second, birds performing R3s had, after their initial error run, learned to avoid both stimuli, and so switched control to irrelevant features (in this case, primarily to position). Further exploration of the within-day serial reversal situation is therefore expected to throw light on the role of inhibitory

processes in the acquisition and reversal of simultaneous discriminations, and so on to the nature of the mechanisms disrupted by certain brain lesions.

The experiment reported here was designed to test an implication of the hypothesis outlined above, namely, that where no salient irrelevant dimension is available, the superiority of R2s over R3s should be markedly reduced. In this experiment, choices were made between two keys, each of which was illuminated by red light of approximately the same intensity, so that no obvious alternative dimension was available to control the subjects' behaviour. The experiment also assessed the generality of the finding that R1s are inferior to R2s by introducing the following differences (besides the change in dimension) from the procedure used in Macphail's (1970) experiment. First, both naive birds and birds with varied prior experience served as subjects; second, incorrect choices were not punished by turning off the houselight, and third, a response requirement was introduced for each choice.

Method

Subjects

The subjects were six pigeons (*Columba livia*) that were maintained throughout at 80% of their *ad lib* feeding weights. Three of the birds had extensive prior experience in a variety of test situations, which included undergoing successive discriminations, generalization tests, and experimental extinction. The other three subjects were experimentally naive.

Apparatus

Animals were tested in a pigeon chamber (30 × 30 × 30 cm), which contained three response keys, each 2 cm in diameter, and was placed in a dark sound-attenuated cubicle. The centre key was mounted over a grain feeder at a height of 20 cm above the floor. The remaining two keys were mounted on either side of the centre key at the same level. The keys were 8 cm apart, centre to centre. The centre key could be illuminated from behind by white light; the side keys could be illuminated by red light. A houselight in the roof of the chamber was continuously illuminated and white noise was also provided throughout.

The control of the sequence of events in the chamber and the collection of data were carried out on-line by an Elliott 4130 computer. Programmes were written in the Experiment Control Language developed in this laboratory (Francis and Sutherland, 1969).

Procedure

Pretraining: Subjects were pretrained to peck all three keys by the use of an auto-shaping procedure (Brown and Jenkins, 1968).

Training: Each trial commenced with the illumination of the centre key: a response to this key extinguished it and illuminated the two side keys with red light. Each peck on a side key produced a clearly audible feedback click from a relay attached to the outside wall of the chamber. A total of five responses to one of the two keys constituted a choice, following which both side keys were extinguished. Where the choice was correct, subjects were given 4 sec access to grain in the hopper, followed by a 10 sec inter-trial interval; incorrect choices were followed by a 14 sec inter-trial interval.

Acquisition and each reversal continued until the criterion of 10 successive correct choices had been achieved, whereupon the reinforcement value of the side keys was reversed; this procedure continued until the three daily runs had been completed, when the subject was removed from the apparatus. The experiment consisted of 15 days on which subjects were given three daily runs to criterion; the right-hand key was correct for the first run of Day 1 (acquisition), and the side of the correct key was reversed overnight throughout the 15 days.

Results

For purposes of analysis, trials were classified as: (a) initial errors (errors which preceded the first correct choice of a reversal); (b) correct choices (excluding the 10 trials of the criterion run of a reversal); and (c) errors, other than initial errors. The classification of individual trials in this experiment is, as a result of the absence of an irrelevant dimension, different in some respects from that used in the Macphail (1970) experiment. In the latter analysis, an attempt was made to distinguish between responding systematically to position and to colour, by classifying responses as position or colour responses only when four responses in succession were to the same side or the same colour (there being no more than three successive trials with a given colour on the same side). In the present experiment, there is, clearly, no corresponding rationale for selecting a run of four as necessary to indicate systematic responding to one side, and the most parsimonious assumption, namely, that all choices are controlled by position cues, has been made. Thus in this experiment the total correct choice score of each reversal corresponds to the correct run score of the previous experiment. It should be noted that this procedure in fact remains close to the former procedure, as, in both analyses, runs of correct choices are weighted according to their length: that is, in the 1970 experiment, the correct run score of a given reversal was the total number of trials scored as members of correct runs, so that a run of eight correct choices was equivalent to two runs of four correct choices. So in this experiment, although the criterion for a "correct run" now stands at one correct response, runs of longer length receive correspondingly more weight in the overall analysis.
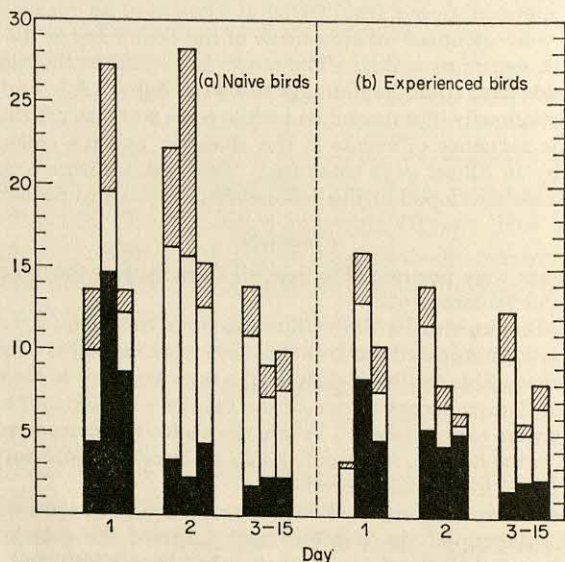
FIGURE 1. The left-hand column of each triple column shows mean R_1 performance (or, on Day 1, acquisition performance), the centre and right-hand columns show mean R_2 and R_3 performance respectively. The third triple column in each part of the figure shows R_1, R_2 and R_3 performance averaged over the last 13 days. ■ Initial errors; □ correct choices; ▨ other errors.

It can be seen from Figure 1 that the experienced subjects enjoyed a considerable advantage over naive birds over the first two days' training, but that the performance of the two groups was comparable over the remainder of the series. In each group, R2s required, over the final 13 days, considerably fewer trials than R1s, and some-what fewer than R3s; the major source of the difference between R1s and R2s lay in the relatively small number of correct choices preceding the criterion run in R2s.

As in the Macphail (1970) experiment, the overall pattern of performance emerged very rapidly; on Day 3, five of the six birds took more trials to criterion on R1 than on R2. Statistical analysis was therefore based on performance on Days 3–15. Mean scores for each group on the various measures were obtained for all three types of reversal on each of the 13 days; differences between reversals were assessed by subjecting these mean scores to the Wilcoxon Matched Pairs Signed-Ranks Test (2-tailed), where $N = 13$. The experimental and naive groups were treated separately.

The principal findings were as follows: for both experienced and naive subjects, R1s required more trials to criterion than R2s ($P < 0.01$ in both cases) or R3s ($P < 0.05$ in both cases); in neither group was there, however, any significant difference between R2s and R3s on this measure. As would be expected from these results, there were more correct choices in R1s than in R2s ($P < 0.01$ in both groups) or R3s ($P < 0.05$ in both cases), and there were no significant differences, in terms of correct choices, between R2s and R3s in either group. When the total number of errors (including initial errors) were compared across reversals, R1s were inferior to R2s for the experienced group ($P < 0.02$), but did not differ significantly for the naive group. There were more total errors in R3s than in R2s for both experienced ($P < 0.01$) and naive subjects ($P < 0.05$). As can be seen from Figure 1, the mean number of initial errors remained very low in all three types of reversal, and differences between reversals did not achieve statistical significance on this measure for either group.

The overall pattern of responding was common to all six birds; for example, all six subjects had more correct choices in R1s than in R2s on more than 50% of days (excluding days on which the scores were equal). Moreover, all six birds had more runs in which no error occurred following the first correct choice of a reversal in R2s than in either R1s or R3s; this latter finding is of some interest, as the difference between R2s and R3s, as measured by correct choices or trials to criterion, was not significant.

Discussion

This experiment demonstrates the robustness of the phenomenon of principal interest—namely, the difficulty found by pigeons in successfully sustaining runs of correct choices in R1s as opposed to R2s. The effect is found in all subjects, whether colour or position is the relevant dimension, and relatively independent of previous experience or of variations in testing conditions.

Given the present interpretation of this phenomenon, an important question arises, namely, why should there be less inhibition generated on the negative key in R1s than in R2s? For, unlike the pattern seen in the 1970 experiment, there

was no significant difference in initial errors between R1s and R2s in this experiment, and yet the large difference in correct choice scores still emerged. Moreover, subjects in R1s made, overall, more errors than in R2s (significantly more, in the case of the experienced group). A solution of this dilemma, which relies on concepts drawn from frustration theory (e.g. Amsel, 1962), is proposed: in the presolution phase of R1s, there is at no stage a strong preference (i.e. a high expectancy of reward) for the negative stimulus, so that little frustration should be generated on account of errors, and the negative stimulus should not become aversive. In the early stages of R2s (and R3s), subjects have a high expectancy of reward on the negative key, so that errors should occasion more frustration, and the negative stimulus should become aversive.

The use of position as the relevant dimension eliminated the superiority of R2s over R3s, in terms of trials to criterion; the difference between R2s and R3s on this measure was highly significant when colour was relevant (Macphail, 1970). However, there were indications that R2s did still enjoy some advantage: first, there were more total errors in R3s than in R2s, and, second, more of the reversals in which no errors occurred following the first correct response were R2s than R3s. It may be concluded, therefore, that, in the absence of a salient irrelevant dimension, the advantage of R2s over R3s is severely attenuated, but not entirely abolished. The fact that there still remained a difference between R2s and R3s in this experiment may be accommodated within the present account by supposing that the subjects may have used some alternative less obvious hypothesis (like, for example, response alternation) in R3s.

In the earlier discussion of within-day serial reversal performance (Macphail, 1970), it was suggested that R1s were, due to the absence of initial preferences, comparable to the original acquisition of a conventional simultaneous discrimination, and evidence to suggest that inhibition might not be generated in conventional acquisition performance was presented. Subsequent experimental work (Macphail, 1972) has indicated that the inhibitory strength of the negative stimulus following acquisition of a simultaneous discrimination is no less than that following reversal of such a discrimination. To maintain the parallel between R1s and conventional acquisition, one would, therefore, have to suppose that inhibition is generated in R1s. The differences between R1s and R2s could still be accounted for, although not without difficulty; for example, on the assumption that it is essential, in order to master a simultaneous discrimination, to learn to avoid the negative stimulus, it could be argued that R1s are at a disadvantage for either (or both) of two reasons. First, because the sequences of negative trials essential to the generation of inhibition are, owing to the absence of initial preferences, more likely to have interspersed positive trials in R1s than in the case in R2s; a difficulty with this proposal is that, in this experiment, there were not in fact significantly more initial errors in R2s than in R1s. Second, because errors are less effective in the generation of inhibition in R1s, owing to the relatively low expectation of reinforcement, so that more negative trials are required; the difficulty here is that the difference between R1s and R2s in terms of total errors was, in this experiment, significant only for the experienced group. As both groups showed significant differences, in terms of trials to criterion and correct choices, between R1s and

R2s, the implication is that total number of errors is not the critical determinant of the response pattern observed. A final major difficulty for the proposal that inhibition is generated to a significant extent in R1s is that an entirely new hypothesis would be required to account for the differences seen between R2s and R3s (differences which were observed both in this experiment, and to a much larger extent, in the 1970 experiment); this is because, on this account, subjects will, following the initial errors at the beginning of an R2, have learned to avoid both stimuli, and so be at the same point as they would be in R3s at the corresponding stage. The hypothesis advanced to explain the 1970 results not only predicts a difference between R2s and R3s, but also that this difference should be attenuated where no salient irrelevant dimension is available; the hypothesis receives, therefore, strong support from the results of both experiments, so that, at present, it is most parsimonious to adhere to it, and to accept the not wholly surprising conclusion that there are in fact important differences between R1s and conventional acquisition performance. Differential reward expectancies in R1s as opposed to acquisition are one possible source of such differences: conventional pretraining presumably leads to a high expectancy of reward on both stimuli at the start of acquisition, whereas, as implied in an earlier part of this Discussion, one effect of serial reversal training may be to produce a reduced expectancy of reward at the start of R1s. According to this account, conventional between-day serial reversal training has (at least) two consequences: first, overnight forgetting occurs, and, second, expectancy of reward on the initial trials of each day declines to a level below that obtaining at the beginning of original acquisition.

References

AMSEL, A. (1962). Frustrative nonreward in partial reinforcement and discrimination learning: some recent history and a theoretical extension. *Psychological Review*, **69**, 306–28.

BROWN, P. L. and JENKINS, H. M. (1968). Auto-shaping of the pigeon's keypeck. *Journal of the Experimental Analysis of Behavior*, **11**, 1–8.

DOUGLAS, R. J. (1967). The hippocampus and behavior. *Psychological Bulletin*, **67**, 416–42.

FRANCIS, J. G. F. and SUTHERLAND, N. S. (1969). A system for controlling animal experiments on-line. In MORAY, N. (Ed.), *On-line Computing for Psychology*. Pp. 43–56. Sheffield: University of Sheffield.

MACKINTOSH, N. J., McGONIGLE, B., HOLGATE, V. and VANDERVER, V. (1968). Factors underlying improvement in serial reversal learning. *Canadian Journal of Psychology*, **22**, 85–95.

MACPHAIL, E. M. (1970). Serial reversal performance in pigeons: role of inhibition. *Learning and Motivation*, **1**, 401–10.

MACPHAIL, E. M. (1971). Hyperstriatal lesions in pigeons: effects on response inhibition, behavioral contrast, and reversal learning. *Journal of Comparative and Physiological Psychology*, **75**, 500–7.

MACPHAIL, E. M. (1972). Inhibition in the acquisition and reversal of simultaneous discriminations. In Boakes, R. A. and Halliday, M. S. (Eds), *Inhibition and Learning*. London: Academic Press. In press.

STETTNER, L. J. and SCHULTZ, W. J. (1967). Brain lesions in birds: effects on discrimination acquisition and reversal. *Science*, **67**, 160–76.

IDENTIFICATION VERSUS DISCRIMINATION OF DISTINCTIVE FEATURES IN SPEECH PERCEPTION

SHEILA BLUMSTEIN

Brown University and Boston Aphasia Research Center

AND

WILLIAM COOPER

Brown University

Dichotically presented CV syllables were presented to 24 normal subjects under two contrasting performance requirements: consonant identification and consonant discrimination. In the identification task, subjects made more errors identifying stimuli distinguished by two as opposed to one distinctive feature. Conversely, in the discrimination task, subjects made more errors for stimuli distinguished by one as opposed to two distinctive features. It was proposed that the results of the identification task reflect the degree of information load in short-term memory. The results of the discrimination task, on the other hand, reflect the degree of perceptual similarity between contrasting stimuli. Both results were accounted for within a distinctive feature framework. Analyses of the individual features which comprised one and two feature distinctions demonstrated the perceptual prominence of the feature [voice] in contrast to the two place features [compact] and [grave].

Introduction

It has been shown that when two sounds are presented simultaneously, a different sound to each ear, one ear functions more efficiently than another. If the dichotic pair consists of language stimuli, right handed subjects typically show a right ear superiority (Kimura, 1961*a*, *b*). If the stimuli are non-language sounds such as music or environmental noises, subjects show a left ear superiority (Kimura, 1964; Curry, 1967; Spellacy, 1970). These results have suggested that the ear preferences demonstrated in the dichotic listening technique reflect (a) the greater efficiency of the contralateral auditory pathways over the ipsilateral pathways in transmitting auditory stimuli and (b) the active involvement of both cerebral hemispheres in the perception of specific types of auditory stimuli, the left hemisphere for verbal stimuli and the right for non-verbal stimuli (Kimura 1961*a*, *b*).

Further dichotic studies have revealed that the right ear superiority for verbal material is not limited to real language. Not only have nonsense syllables, backwards speech, and synthetic CV stimuli produced right ear advantages, but also some of the components of the speech signal—consonants and vowels—have produced a right ear effect. Consonants have shown a clear-cut right ear preference

under many conditions (Shankweiler and Studdert-Kennedy, 1966, 1967a, b; Spellacy and Blumstein, 1970; Studdert-Kennedy and Shankweiler, 1970) and although it has been more difficult to show lateralized ear advantages for vowels, significant right ear advantages have been obtained when expectation or set has been manipulated (Spellacy and Blumstein, 1970; Darwin, 1971).

In a recent study, Studdert-Kennedy and Shankweiler (1970) have investigated the effects of dichotic listening on the perception of components of the speech signal. Specifically, they demonstrated significant right ear advantages for initial and final stop consonants and a non-significant right ear advantage for medial vowels. Further analysis of the data revealed that not only were the consonant phonemes lateralized but that the articulatory features of which these consonants were comprised also seemed to be processed independently in the language domin-ant hemisphere. They noted that the features *voice* and *place* were independently affected under dichotic listening conditions. Thus, competing stimuli varying in the feature *place* (e.g. p–t) produced more errors than stimuli that varied in the feature *voice* (e.g. p–b). Similar evidence has been obtained in studies of speech perception under conditions of filtering and white noise (Miller and Nicely, 1955; Singh, 1966, 1969), in short-term memory studies (Conrad, 1964; Klatt, 1968; Wickelgren, 1965, 1966) and in a study requiring subjective judgements of phonemic similarity (Greenberg and Jenkins, 1964).

The notion that sub-phonemic or distinctive features are the basic primes upon which the phonemic systems of language are comprised has been a central issue in phonological linguistic theory since the 1920s (Trubetzkoy, 1964: Jakobson, Halle and Fant, 1961; Jakobson and Halle, 1957; Chomsky and Halle, 1968). Within any language system, every phoneme is characterized by the minimum number of features needed to distinguish it from all other phonemes of that system. Implicit in the theory is the notion that phonemes distinguished by one distinctive feature are more closely aligned structurally, articulatorily, and acoustically, as well as psychologically, than phonemes separated by several features (cf. Jakobson, 1968). Thus, for example, it is theorized that p–b, distinguished by the feature *voice*, and p–t, distinguished by the feature *place*, are more similar than are p–d, dis-tinguished by two features—*voice* and *place*.

The relationship between phonemes distinguished by one and two features was considered briefly in the Studdert-Kennedy and Shankweiler study. They found that when a phoneme pair contrasted in only one feature, subject errors were less likely to occur than when a phoneme pair contrasted in two features. They concluded that the duplication of auditory information resulted in the observed advantage for competing stimuli differing by one distinctive feature. Moreover, in a later study, Studdert-Kennedy et al. (1971) demonstrated that this advantage was based on phonetic similarity rather than on auditory or acoustic similarity.

These results have important implications for the nature of the speech percep-tion process. The observed superiority in the identification of stimuli distinguished by one in comparison to two distinctive features is compatible with the notion that each distinctive feature conveys an independent but differing amount of information. Furthermore, it suggests that the difficulty the subject has in identifying competing stimuli is a function of the number of distinctive features which the subject must

extract, hold in short-term memory, and subsequently encode. Here, the assumption is that a subject must process a greater amount of information to distinguish stimuli differing by two features. However, although the identification task may reflect the nature of the information loss which can occur in the context of competing stimuli, it does not reflect the nature of the bond existing between the competing stimuli. Subjects may have more difficulty identifying stimuli two features apart because a greater amount of information has to be processed. But since such stimuli are further apart acoustically, articulatorily, and perceptually, they should also be judged as more different than those only one feature apart.

It is the object of this study to explore the relationship between competing stimuli in the dichotic listening situation. It is hypothesized that subjects' performance in the perception of competing stimuli distinguished by one versus two distinctive features will produce opposite effects depending upon whether the task tests short-term memory or discrimination. Specifically, it is proposed that if a subject is required to judge whether competing stimuli are the same or different (discrimination test), there will be more errors on stimuli distinguished by one distinctive feature in comparison to two distinctive features. Conversely, if subjects are required to identify the competing stimuli (identification test), more errors should occur on stimuli distinguished by two as opposed to one distinctive feature.

Method

Stimulus materials

The test stimuli used in this experiment were CV syllables differing only in their initial consonants. The consonants consisted of the English stop series /p t k b d g/ paired with the vowel /a/. The stimuli were pronounced by a trained linguist at an even intensity and uniform pitch contour. The syllables were recorded and subsequently edited using the pulse code modulation system made available at the Haskins Laboratory (Cooper and Mattingly, 1969). With this system, the initial stop consonants were synchronized for onset with an accuracy of 500 μsec and the over-all intensity of each CV syllable was equalized. The duration of each stimulus varied from 350–375 msec.

The test items were paired so that there was an equal occurrence of each distinctive feature used to differentiate the stop consonant series (Jakobson, Halle and Fant, 1961). Thus, five classes of distinctive features were used. These included the feature [voice] used to distinguish the voiceless consonants (e.g. p t k) from their voiced counterparts (e.g. b d g); two features of *place*, the first [grave] used to distinguish peripherally produced consonants from centrally produced consonants (e.g. p–t, b–d) and the second [compact] used to distinguish consonants produced in the front part of the mouth from those produced in the back of the mouth (e.g. p–k, t–k, b–g, d–g); and finally, two classes of distinctive features used to differentiate consonants which share neither the voicing nor place features—[voice/grave] (e.g. p–d, b–t) and [voice/compact] (e.g. p–g, t–g, b–k, d–k).

The identification task involved two randomized series of dichotic pairs. The first series included four occurrences of each distinctive feature, the second series consisted of the same dichotic pairs recorded in a different order with channels reversed. Thus there were a total of 40 test items and eight occurrences of each distinctive feature.

The discrimination task involved a total of 80 dichotic pairs. The first 40 were the same test items used in the identification task. The second 40 consisted of the same items as in the first part but presented in a different order and recorded with channels reversed. Thus in this task there were 16 occurrences of each class of distinctive feature.

Subjects

There were a total of 24 subjects, 12 male and 12 female who were primarily graduate and undergraduate students at Harvard University. All reported that they were strongly right-handed and had no known hearing impairment.

Procedures

Subjects were tested individually in a quiet room on a Tandberg 3000X tape recorder. The output of Koss Pro 4AA earphones was matched at 70 dB by means of a calibration tone (1,000 Hz) measured on a General Radio sound level meter (type 1565 A).

All subjects took both the identification and discrimination task. One half received the identification test first and the other half received the discrimination test first. Within each task, earphones were reversed for half the subjects.

In the identification task, subjects were told that they would hear two different simultaneous stimuli, both taken from the series pa, ta, ka, ba, da, ga. They were instructed to write down what they heard in any order they wished but to always give two different responses even if they had to guess. Subjects wrote their responses on answer sheets provided. At the top of each sheet appeared the six CV syllable test items.

On the discrimination task, subjects were told that they would hear two simultaneous stimuli and were instructed to judge whether they were the *same* or *different* to each ear. They were not told that these stimuli were always different. Subjects had only to check on an answer sheet in the appropriate column whether these stimuli were the same or different.

Preceding each task, subjects heard the six stimuli binaurally. They then received three practice examples for each particular task. In both tasks, there was a 5-sec interval between trials. A short rest period was given to each subject between the administration of the two tasks.

Analysis of Results

The subjects' performance on the identification versus discrimination of one and two distinctive features is illustrated in Figure 1. A two-tailed *t*-test revealed that subjects had significantly greater difficulty identifying stimuli distinguished
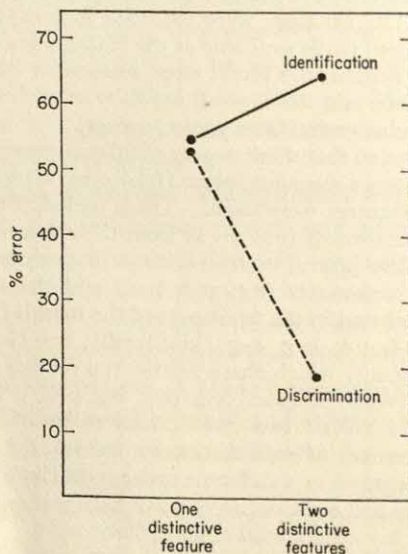
FIGURE 1.   Mean per cent error on one and two distinctive features in the report and discrimination tasks.

by two distinctive features in comparison to those distinguished by one distinctive feature ($t = 2.54$, $df = 1$, $P < 0.02$). Conversely, in the discrimination task, they had significantly greater difficulty discriminating stimuli distinguished by one in comparison to two distinctive features ($t = 11.36$, $df = 1$, $P < 0.001$). Moreover, the subjects' performance on stimuli distinguished by one distinctive feature did not differ significantly in the discrimination and identification tasks ($t = 0.48$, $df = 1$, $P > 0.20$), whereas a comparison of the subjects' performance on stimuli distinguished by two distinctive features on the two tasks showed that they performed significantly better on the discrimination task ($t = 13.74$, $df = 1$, $P < 0.001$).

The one versus two distinctive feature categories were analysed into their component features to determine whether any particular feature or features contributed to the ease or difficulty of identification, on the one hand, and discrimination, on the
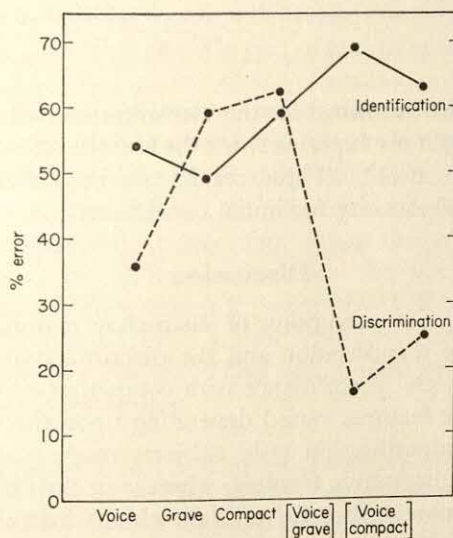
FIGURE 2. Mean per cent error on the individual distinctive features in the report and discrimination tasks.

other. The results are illustrated in Figure 2. The data were tested for significance by means of an analysis of variance and Scheffé test. A comparison of the individual features in the two tasks revealed that the features [voice], [voice/compact], [voice/grave] were significantly easier for the subjects in the discrimination task than in the identification task ($t = 3.33$, $df = 1$, $P < 0.01$; $t = 7.00$, $df = 1$, $P < 0.01$; and $t = 9.43$, $df = 1$, $P < 0.01$, respectively). Within the identification task, only the features [grave] and [voice/grave] were significantly different from each other ($t = 3.63$, $df = 4$, $P < 0.05$). For the discrimination task, a comparison of the individual features revealed significant differences between a number of them, and these results are shown in Table 1.

TABLE I

Feature analysis of discrimination task

Features	F-value*	Significance
[voice]–[compact]	4·84	$P < 0·01$
[voice]–[grave]	4·55	$N < 0·01$
[voice]–[voice/compact]	2·00	n.s.
[voice]–[voice/grave]	3·26	$P < 0·05$
[compact]–[grave]	0·30	n.s.
[compact]–[voice/compact]	6·85	$P < 0·01$
[compact]–[voice/grave]	8·10	$P < 0·01$
[grave]–[voice/compact]	6·55	$P < 0·01$
[grave]–[voice/grave]	7·81	$P < 0·01$
[voice/compact]–[voice/grave]	1·26	n.s.

*$df = 4$

Ear preferences were determined for the identification task. A two-tailed t-test revealed a significant right ear superiority for the identification of competing stimuli ($t = 3·62$, $df = 1$, $P < 0·01$). These results are consistent with the previous reports of a right ear superiority for initial consonants.

Discussion

This study explored the perception of distinctive features in relation to two experimental tasks, the identification and the discrimination of dichotic stimuli. It was shown that subjects' performance with competing stimuli distinguished by one and two distinctive features varied depending upon the nature of the experimental task. In the identification task, subjects made more errors for stimuli which differed by two distinctive features, whereas in the discrimination task, the same subjects made more errors for stimuli which differed by one distinctive feature.

The experimental tasks tap two different but related aspects of speech perception. In the identification task, the subject must analyse the auditory information into its linguistic components and hold it in short-term memory long enough to be able to encode his response. For those stimuli distinguished by one distinctive feature, all feature values are shared, apart from either place or voice. Those stimuli distinguished by two distinctive features, on the other hand, share neither place nor voice. As a result of the greater shared information in the case of one distinctive feature differences, there is less total information which must be extracted and held in order to identify competing stimuli. Consequently, fewer errors are made in identifying stimuli distinguished by one as opposed to two features.

In the discrimination task, however, the storage of information is not a factor. Instead, the subject need only judge the relative similarity between competing stimuli. Because those stimuli perceptually further apart are easier to distinguish from each other than those which are more similar, the subject is able to discriminate

two feature distinctions more readily than one feature distinctions. Thus, unlike the identification task, the discrimination task reveals more errors for one feature distinctions.

As Figure 1 demonstrates, the crucial difference between the subjects' performance on both tasks is with the two distinctive feature differences. Here, discrimination is significantly easier than identification. Performance with one distinctive feature differences, on the other hand, does not differ significantly in either the identification or discrimination task, i.e. it is just as easy for the subject to identify as to discriminate these stimuli. These results suggest that the major factor contributing to the ease or difficulty of each task is the addition of another feature contrast between competing stimuli.

The analyses of the individual feature values which comprise the one and two distinctive feature differences indicate that not all distinctive feature values are perceived equally well. Perhaps the most consistent finding was the prominence of the feature [voice]. The feature analysis of errors comparing the two tasks revealed that discrimination is significantly easier than identification whenever (and only whenever) a voice distinction is part of the difference between competing stimuli (see Fig. 2).

Further evidence for the prominence of voicing is provided by a feature analysis of errors within each task. Within the identification task, there is a small superiority in reporting stimuli distinguished by voicing than for stimuli distinguished by place (54·22% vs. 53·12%). Within the discrimination task, the voicing distinction is detected significantly better than either the grave or compact place distinctions. Moreover, the difference between the discrimination of stimuli distinguished by a *place* feature versus a [voice/place] distinction is significantly greater than the difference between stimuli distinguished by a *voice* versus a [voice/place] distinction (see Fig. 2). This again suggests the prominence of the feature [voice] in speech perception.

References

CHOMSKY, N. A. and HALLE, M. (1968). *The Sound Pattern of English*. New York: Harper & Row.

CONRAD, R. (1964). Acoustic confusions in immediate memory. *British Journal of Psychology*, 55, 75–84.

COOPER, F. S. and MATTINGLY, I. G. (1969). Computer-controlled PCM system for the investigation of dichotic speech perception. Paper presented at the 79th meeting of the Acoustical Society of America, Philadelphia.

CURRY, F. K. W. (1967). A comparison of left-handed and right-handed subjects on verbal and non-verbal dichotic listening tasks. *Cortex*, 3, 343–52.

DARWIN, C. J. (1971). Ear differences in the recall of fricatives and vowels. *Quarterly Journal of Experimental Psychology*, 23, 46–62.

GREENBERG, J. H. and JENKINS, J. J. (1964). Studies in the psychological correlates to the sound system of American English. *Word*, 20, 157–77.

JAKOBSON, R. (1968). *Child Language, Aphasia, and Phonological Universals*. (Trans. A. R. KEILER.) The Hague: Mouton.

JAKOBSON, R. and HALLE, M. (1957). *Fundamentals of Language*. The Hague: Mouton.

JAKOBSON, R., HALLE, M. and FANT, G. (1961). *Preliminaries to Speech Analysis*. Cambridge, Mass.: M.I.T. Press.

KIMURA, D. (1961a). Some effects of temporal lobe damage on auditory perception. *Canadian Journal of Psychology*, **25**, 156–65.

KIMUERA, D. (1961b). Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology*, **15**, 166–71.

KIMURA, D. (1964). Left–right differences in the perception of melodies. *Quarterly Journal of Psychology*, **16**, 355–8.

KLATT, D. H. (1968). Structure of confusions in short-term memory between English consonants. *Journal of the Acoustical Society of America*, **44**, 401–7.

MILLER, G. and NICELY, P. E. (1955). An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America*, **27**, 338–52.

SHANKWEILER, D. and STUDDERT-KENNEDY, M. (1966). Lateral differences in perception of dichotically presented synthetic consonant-vowel syllables and steady-state vowels. *Journal of the Acoustical Society of America*, **39**, 1256(A).

SHANKWEILER, D. and STUDDERT-KENNEDY, M. (1967a). An analysis of perceptual confusions in identification of dichotically presented CVC syllables. *Journal of the Acoustical Society of America*, **41**, 1581 (A).

SHANKWEILER, D. and STUDDERT-KENNEDY, M. (1967b). Identification of consonants and vowels presented to left and right ears. *Quarterly Journal of Experimental Psychology*, **19**, 59–63.

SINGH, S. (1966). Cross-language study of perceptual confusions of plosive phonemes in two conditions of distortion. *Journal of the Acoustical Society of America*, **40**, 635–56.

SINGH, S. (1969). Interrelationship of English consonants. *Proceedings of the International Congress of Phonetic Sciences*, 6th, Prague, 542–4.

SPELLACY, F. (1970). Lateral preferences in the identification of patterned stimuli. *Journal of the Acoustical Society of America*, **47**, 574–8.

SPELLACY, F. and BLUMSTEIN, S. (1970). The influence of language set on ear preference in phoneme recognition. *Cortex*, **6**, 430–9.

STUDDERT-KENNEDY, M. and SHANKWEILER, D. (1970). Hemispheric specialization for speech perception. *Journal of the Acoustical Society of America*, **48**, 579–94.

STUDDERT-KENNEDY, M., SHANKWEILER, D. and PISONI, D. (1971). Auditory and phonetic processes in speech perception: evidence from a dichotic study. Presented at the 81st Meeting of the Acoustical Society of America, Washington D.C.

TRUBETZKOY, N. S. (1964). *Principes de Phonologie*. Paris: Librairie C. Klincksieck.

WICKELGREN, W. (1965). Acoustic similarity and intrusion errors in short-term memory. *Journal of Experimental Psychology*, **70**, 102–8.

WICKELGREN, W. (1966). Distinctive features and errors in short-term memory for English consonants. *Journal of the Acoustical Society of America*, **39**, 388–98.

INHIBITION OF *AD LIBITUM* FEEDING IN RATS BY SALT INJECTIONS AND WATER DEPRIVATION

F. M. TOATES AND KEITH OATLEY

Laboratory of Experimental Psychology, University of Sussex

Inhibition of *ad libitum* feeding in rats was induced by hypertonic NaCl injections. Though osmotic loads of sufficient size were capable of abolishing feeding completely for a time, the effect was not as large as had been predicted from a hypothesis of strictly linear subtractive inhibition. Feeding at a low level of hunger seems to be somewhat less affected by osmotic inhibition than feeding on a deprivation schedule. Inhibition of feeding was also produced by deprivation of water, and both the inhibition of food intake during deprivation, and the disinhibition by subsequent drinking indicated that the amount of inhibition of food intake is a non-linear (accelerating) function of water deficit. A model of the process indicating that the thirst signal undergoes a non-linear transformation before being subtracted from the signal corresponding to food demand is proposed.

Introduction

Reduction of feeding by thirsty animals is probably due to the direct action of the system which controls water balance, on the system controlling feeding (Oatley, 1967). The implication of models of the control of drinking, constructed by McFarland (1965) and Oatley (1967), is that this inhibition of eating by thirst is a subtractive process, such that a particular level of thirst diminishes the amount eaten by a hungry animal by some particular quantity. In Barbary doves McFarland (1964) found that deprivation of water for 96 hr produced a secondary deficit equivalent to 12 hr of food deprivation. More recently Oatley and Toates (1971) have estimated an inhibition factor (the amount by which eating was reduced by thirst in rats) and found that irrespective of the level of hunger, an osmotic stimulus requiring 1 ml of water to be drunk, reduced food intake by about one-third of a gramme.

McFarland's (1965) and Oatley's (1967) models also predict disinhibition of feeding by drinking, and Kakolewski and Deaux (1970) found precisely this in experiments where following deprivation of water the drinking of 3 ml of water initiated eating following a latency. When solutes were added to the drinking water the latency increased as a function of the solute concentration. Addition of solutes also has the effect of reducing the size of a subsequent meal (Hsiao and Kendrick, 1970; Hsiao and Langenes, 1971). On the basis of these experiments it would seem that food intake is restrained by the osmotic concentration of the body fluids, and reducing osmotic pressure results in a disinhibition which allows a certain amount of food to be eaten. This conclusion was also echoed in the experiments of Oatley

and Toates (1971) who calculated the effect of drinking on eating, and found the value of the disinhibition factor to be very similar to that displayed during inhibition, namely 1 ml of water was worth about one-third of a gramme of food.

The experiments which follow are further tests of the hypothesis that inhibition occurs subtractively; the experiments investigate whether thirst reduces feeding by a specific quantity independently of the level of hunger; and whether the inhibition factor is a constant. In Experiment I we investigated whether osmotic injections could completely abolish feeding, as predicted by a subtractive hypothesis.

Experiment I

It has been found (Oatley and Toates, 1971) that an intraperitoneal injection of 2 ml of 8% NaCl reduced eating by food-deprived rats by 2·9 g. The hypothesis of subtractive inhibition predicts that rats which would eat less than 2·9 g without an osmotic load, would, with the injection of 2 ml of 8% NaCl, have their eating completely stopped. In this experiment, therefore, rats fed *ad lib*, whose intake was relatively low, were given similar treatments to those of the earlier experiment. There was some suggestion in the previous study (Oatley and Toates, 1971) that the disturbance caused by repeated weighing of food provoked eating. Therefore in the present experiment we compared how much rats ate over a period of time when disturbed on several occasions with the amount eaten over the same period of time when the animals were undisturbed.

Procedure

Subjects were six male hooded rats of mean weight 286 g at the beginning of the experiment. They were housed in individual metabolism cages (described in detail elsewhere: Oatley and Tonge, 1969). Food was Spiller's Small Animal Diet, in powdered form, available from a food compartment at one end of the cage. Tap water was available from a burette at the other end. An artificial 12-hr dark-12-hr light cycle was employed, lights coming on at 08.00 hr each day.

All animals received one habituation injection of 2 ml of 1-M NaCl intraperitoneally two days before the experimental tests began.

The first part of the experiment (the undisturbed eating test) consisted of giving half of the rats an intraperitoneal injection of 2 ml of 0·9% saline, and the other half, 2 ml of 8% saline. Water was removed at the time of the injection but food was available *ad lib* throughout. After the injections, the amounts of food eaten were measured to the nearest 0·1 g over a period of 7 hr (09.00–16.00 hr) during which the animals were left undisturbed. At 16.00 hr, water was again made available and two days of recovery followed before the main part of the experiment.

On each test day in the main section of the experiment, water was removed at 09.00 hr and the animals were given one of six treatments: injection of 2 ml of 0·9, 4 or 8% NaCl (i.p.), together with a ration of either 3 ml water at 4 and 5½ hr after the injection or no water throughout the test period until 16.00 hr. All six treatments were presented to all animals in a counterbalanced order. The 1-hr period following the injection corresponded to an hour's recovery from the injection, when, in the earlier experiment (Oatley and Toates, 1971) no food was allowed. But for the present experiment, the *ad lib* food schedule necessitated the presence of food. Measurements of food intake were made at 1, 1½, 2, 3 and 4 hr after the injection, exactly as in the earlier experiment. At 13.00 hr either 3 ml of water or no water was presented, and from then on measurements were made of food eaten and water drunk every half hour in order to determine the effect of drinking. Animals

receiving water at 13.00 obtained another 3 ml at 14.30. Water at those times was offered in a small cup attached to the side of the cage, the animals previously having been given experience of drinking from these cups. At the end of the test the animals were again allowed water *ad lib.* Six test days were needed to allow all animals each treatment once; and each test day was separated by one day of recovery.

Two days after this part of the experiment, those subjects which received 0·9% NaCl on the first experimental day were injected with 8% NaCl and *vice versa.* A 7-hr undisturbed eating test then followed as before.

Results

The results are shown in Figure 1. During the first hour the 8% injection reduced eating to virtually zero, a mean of 0·06 g being eaten. Subsequently, however, animals did begin to eat on trials following 8% NaCl.
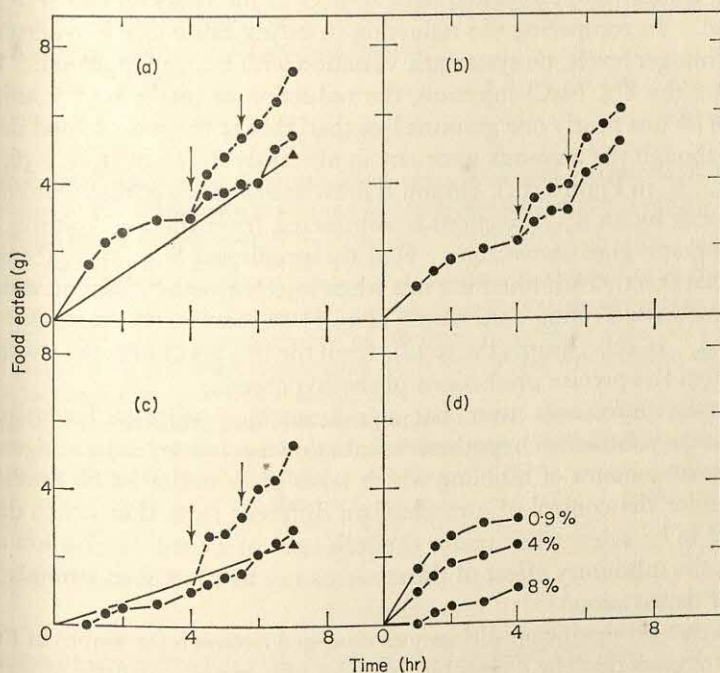
FIGURE 1. Food eaten following various hypertonic saline injections, (a) 0·9%, (b) 4% and (c) 8%. ●——● food eaten without water, ▲ 7-hr score when the animals were undisturbed, ●---● food eaten when water was available. The arrows mark the times at which 3 ml were given. Part (d) enables the various injections to be compared up to the point where water was made available. Beyond this time the density of points makes individual graphs essential.

The differences between the 0·9 and 4% injected groups were at no time significant at the 0·05 level. At all observations the differences between the 4 and 8% groups were significant at the 0·05 level ($2·9 < t < 3·75$, $df = 5$).

In analysing the effects of giving water from 13.00 onwards, regression lines were fitted to the data for the amount of food eaten over each 1½-hr observation period as a function of the amount of water drunk in this period. No significant differences were found at the 0·5 level between the slopes of the regression lines for the various concentrations injected nor between the first and second presentations of water.

The amount of food eaten without water over the 7-hr period, when the animals were not disturbed, was no different from the total eaten over the same period when they were disturbed ten times (0·9% injection, $t = 0·81$, $df = 5$, $P > 0·4$; 8% injection, $t = 0·88$, $df = 5$, $P > 0·4$).

Discussion

As Figure 1 shows, 2 hr after injection, the reduction in feeding produced by the 8% injection was 2·1 g. This may be compared to 3·0 g for animals on a schedule maintaining them at 98% of their growing body weight, and 2·9 g for animals kept at 85% of their growing body weight, in earlier experiments (Oatley and Toates, 1971). All values were calculated with respect to the trials on which 0·9% saline was injected. In comparing the reduction in eating caused by a hypertonic injection across hunger levels, no systematic variation with hunger appeared. However, by 2 hr after the 8% NaCl injection, the reduction of intake by the animals that were fed ad lib was nearly one gramme less than that of the rats on food deprivation schedules, though the amounts were not significantly different ($t < 1·36$, $df = 10$, $P > 0·05$). If, in Figure 1(d), a graph is drawn such that 3 g (the inhibition caused in hungry rats by an 8% injection) is subtracted from the 0·9% score, it is 3 hr before this graph goes above zero. Had the prediction from the subtractive hypothesis been exactly confirmed the rats when injected with 8% saline would therefore not have started eating until hour 3 (Fig. 1) whereas by that time they had consumed 0·6 g. In other words the results from the 8% NaCl injections were slightly different from the precise predictions of the hypothesis.

Though the differences were not significant they were sufficiently large to suggest that the subtractive hypothesis breaks down at low levels of hunger. There may be a small amount of nibbling which takes place under ad lib conditions and which is under the control of a mechanism different from that which determines the amount to be eaten when under the influence of a food deprivation schedule. If so, then the inhibitory effect of thirst seems not to affect it so strongly as eating induced by deprivation.

The fact that no significant differences emerged between the slopes of the regression lines suggests that the disinhibitory effect of a certain quantity of water allows a particular quantity of additional food to be eaten, and that within the range of injections employed this factor was independent of the level of thirst. Mean slopes of the regression lines (measured in g/ml) in the present experiment were 0·40 (0·9% injection) 0·20 (4% injection) and 0·39 (8% injection). Thus, as in the earlier experiments (Oatley and Toates, 1971), each 1 ml of water drunk was worth about one third of a gramme of food (averaged across all conditions).

It is a common observation that merely disturbing an animal which had food and water ad lib will result, shortly afterwards, in the animal eating, drinking or doing both (Valenstein, Cox and Kakolewski, 1970; Oatley, 1971). The question that arises in connection with this phenomenon is whether or not the animal eats and drinks in excess of need under these conditions. Comparison of the total food intake over the 7-hr eating period following 0·9% and 8% saline injections in the main part of the experiment where rats were disturbed 10 times by food weigh-

ings, with the undisturbed trials before and after the main series, suggests that over the long term the amount eaten was unaffected by whether the animal was disturbed or not, although disturbances may possibly have had the effect of bringing eating forward in time (see Fig. 1(a)).

Experiment II

After a hypertonic salt injection, the amount of disinhibition of eating caused by a fixed amount of water has been found not to be dependent upon the level of hunger or hypertonically induced thirst. In this experiment thirst induced by deprivation of water was investigated, and rats which received food *ad lib* were given either 24 or 48 hr of water deprivation and then allowed 8 ml of water, in order to see how much disinhibition of feeding resulted.

Procedure

Eight male hooded rats with a mean weight of 269 g took part. They lived throughout in the individual metabolism cages and under the same conditions as in Experiment I.

A 2 × 2 factorial design was employed, with the subjects being deprived of water but not food for 24 or 48 hr: deprivation was followed by a 3-hr test during which food was still available *ad lib* and either 0 or 8 ml of water was allowed for drinking. At the beginning of an eating test water was presented in small cups that were fixed to the side of the cage. The rats had previously been given experience of drinking from these cups, and during the experiment every rat offered water drank all of the 8 ml. Food was weighed to the nearest 0·1 g at the beginning and end of the eating tests.

Repeated measures were taken until all eight rats had received each of the four treatment combinations in a counter-balanced order. Between deprivation sessions, at least three clear days with water *ad lib* were allowed.

Results

TABLE I

Food eaten in 3-hr test session following either 24 or 48 hr water deprivation, and with either 0 or 8 ml of water available

	0 ml	8 ml
24 hr	1·0	2·63
48 hr	0·1	3·04

Table I shows the results of the experiment. The mean difference between the amount of food eaten at 3 hr, with and without water was 1·63 g following 24-hr deprivation of water, and 2·94 g after 48-hr deprivation. By considering food eaten with water, minus food eaten without water, as being the amount of eating attributable to disinhibition, the 48-hr deprived subjects obtained significantly more disinhibition from 8 ml of water than those deprived for 24 hr ($t = 3\cdot14$, $df = 7$,

$P < 0.01$). There was a significant difference between the amount eaten with water and the amount eaten without water for both the 24 hr ($t = 4.65$, $df = 7$, $P < 0.01$) and 48 hr ($t = 9.18$, $df = 7$, $P < 0.005$) treatments.

Discussion

The disparity between the disinhibition following 24 hr and 48 hr seems to have been caused by the fact that after 24 hr of water deprivation eating had not completely stopped, so that rats managed to eat 1 g of food even when not allowed water, whereas after 48 hr of water deprivation, rats ceased eating almost completely. The difference between the actual amounts of food eaten in 3 hr, after 24 hr and 48 hr treatments when receiving water, was not significant ($t = 0.74$, $df = 7$, $P > 0.5$). However, the fixed amount of water in this experiment was more effective for rats deprived of water for 48 hr than for 24 hr, in terms of the extra food it allowed them to eat. Evidently, during an inhibition of eating provoked by thirst, a given amount of water does not allow a fixed extra amount of food to be eaten in all circumstances. This inequality of the disinhibitory effect of 8 ml of water is unlikely to be due to the differences in latent hunger levels, since one of the most consistent findings in the experiments of Oatley and Toates (1971) and, to a lesser extent, in Experiment I of the present paper is that disinhibition is not a function of the level of hunger. Two alternatives remain. As suggested in Experiment I departure from a strictly subtractive process may occur at a low level of food intake; i.e. the possibility exists that after the 24 hr of water deprivation, the 1 g of food eaten in 3 hr was a residual quantity ascribable to a low level of nibbling, under the control of some separate mechanism. After 48 hr of water deprivation, eating without water was reduced effectively to nothing, indicating that thirst of sufficient magnitude can inhibit eating entirely. A more probable explanation of the difference of disinhibitory effects of 8 ml of water for animals deprived of water for 24 or 48 hr, is that during the second day of a two-day deprivation interval, each millilitre of deficit inhibits a larger amount of feeding than during the first day. That is to say, although inhibition may be independent of the level of hunger, it might not be independent of thirst. At higher levels of water deficit the value of the inhibition factor may not be constant, but increasing with the level of thirst.

Experiment III

In this experiment a comparison was made between food eaten with and without water over a 48-hr period in order to see how soon the effect of water deprivation becomes apparent, and whether, as Experiment II suggests, the inhibition of eating increases with increasing deprivation of water. In addition the influence of the circadian rhythm of intake was investigated.

Procedure

Subjects were 16 male hooded rats of mean weight 293 g. They were housed throughout in the same metabolism cages as in Experiments I and II. Other conditions were also similar except in the 12-hr-dark/12-hr-light cycle lights came on at 09.00. There were

two groups of rats, and for one of these (group M), water deprivation started in the morning, at 09.00 hr, while the other (group E) started in the evening at 21.00.   On the first trial half the rats in each group were allowed food and water *ad lib*, while the other half were deprived of water for 48 hr.   On the second trial, after an interval of 4 days with water *ad lib*, the treatments were reversed.   Food, which was available *ad lib* throughout, was measured every 12 hr, at 09.00 hr and 21.00 hr.

Results

Figure 2 shows the results of the experiment.   For group M, starting deprivation at 09.00 hr, no inhibition was apparent for the first 12 hr.   In fact rats without water ate slightly though not significantly more food than when they had water $(t = 1.15, df = 7, P > 0.2)$.   After this point the influence of water deprivation
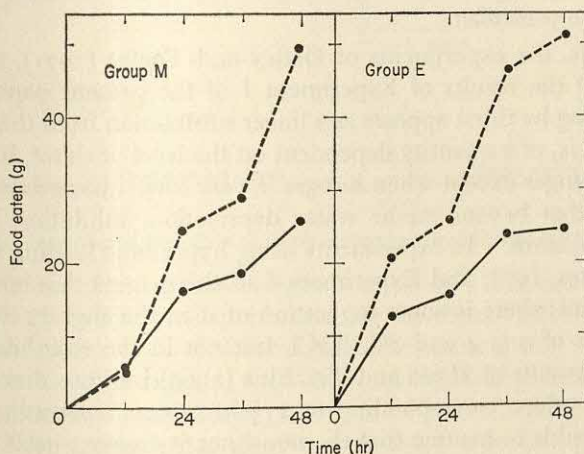
FIGURE 2.   Food eaten by rats with (●– – –●) and without (●——●) water.   Group M started water deprivation at 09.00 hr and Group E at 21.00 hr.

became apparent.   The strong diurnal influence may also be seen for both groups, both with and without water.   Differences between water and no-water treatments were significant at all points $(5.312 < t < 29.34, df = 7, 0.01 > P > 0.001)$, except for group M at 12 hr.

Discussion

At 24 and 48 hr, corresponding mean cumulative eating scores for groups M and E did not differ by more than 1.5 g at any point, and so may be averaged. The mean differences in food intake between trials on which water was, and was not, allowed (i.e. the amounts of inhibition caused by water deprivation) were: 9.29 after 24 hr and 24.9 g after 48 hr.   That is to say, not only did the curves for trials when water was allowed diverge from those on which it was not, but they diverged more sharply on the second day than on the first.   The amount of inhibition produced on the second day of deprivation was 6.5 g more than on the first day, an increase of about 70%.   This is what would be expected on the basis of the finding of Experiment II, that the amount of disinhibition caused by 8 ml of water was 80% more in rats deprived of water for 48 hr than in rats deprived for 24 hr, though

it contradicts the assumption made by Oatley and Toates (1971) that the inhibition factor is a constant, independent of thirst.

General Discussion

As far as consummatory measures (e.g. Stellar and Hill, 1952; Oatley and Dickinson, 1970) or instrumental measures (e.g. Oatley, 1965) have been made, thirst increases as a negatively accelerated function of the interval of water deprivation. Various indices of water balance, e.g. osmolality (Kutscher, 1966), electrolyte concentration (Novin, 1962), increase linearly with deprivation. It seems, therefore, that the positively accelerating inhibition of food intake as a function of period of water deprivation arises in the pathway from the thirst control system to the food intake controller.

In other words, the experiments of Oatley and Toates (1971), and (somewhat more tentatively) the results of Experiment I of the present paper indicate that inhibition of eating by thirst appears as a linear subtraction from the amount eaten; subtraction, that is, of a quantity dependent on the level of thirst, but independent of the level of hunger except when hunger is very low. Experiments II and III, however, show that beyond 24-hr water deprivation, inhibition is a non-linear function of deprivation. In experiments using hypertonic loading to induce thirst (Oatley and Toates, 1971, and Experiment I of this paper) this non-linearity was not clearly evident: there is some suggestion of it in the slightly disproportionate inhibitory effects of 0·9, 4 and 8% NaCl, but not in the disinhibitory effects of drinking. The results of Hsiao and Trankina (1969) indicate that the function is linear for thirst values corresponding to $21\frac{1}{2}$-hr water deprivation and less. It seems most plausible to assume that the non-linearity asserts itself most clearly at higher levels of thirst such as are produced by 48 hr of water deprivation, but not by 2 ml of 8% NaCl (which in rats of approximately this weight produces a thirst equivalent, on several measures, to slightly less than 24 hr of water deprivation (Oatley, 1965)). For lower levels of thirst the equation of $\frac{1}{3}$ g of food with 1 ml of water deficit remains a useful approximation. A less plausible alternative explanation of the differences between this and the earlier experiments would be to assume that the differences between thirst induced by deprivation and by hypertonic loads which reveal themselves in various circumstances (e.g. Grimsley, 1965, Wayner and Petraitis, 1967) are responsible for the non-linear function of deprivation, although the evidence for non-linearity in salt injection experiments was not so striking.

In any case we may represent the form of the inhibitory process as in Figure 3, where the postulates as to the nature of the inhibition of eating during thirst are represented in the formal notation of a control system block diagram compatible with the simulation of thirst of Toates and Oatley (1970). A food demand signal, perhaps arising as a deviation of some index of energy stores from a set point gives rise to eating, but an inhibitory signal which arises from the thirst system can subtract from this signal to reduce the amount eaten or even bring it to zero. The thirst signal is the sum of cellular and extracellular deficits, minus inhibitory effects, and indeed it is known that deficits in these two compartments are additive

in their inhibitory effect on feeding (Hsaio, 1970).   An interaction of this kind causes thirst to subtract a fixed quantity of food, independent of the level of hunger,
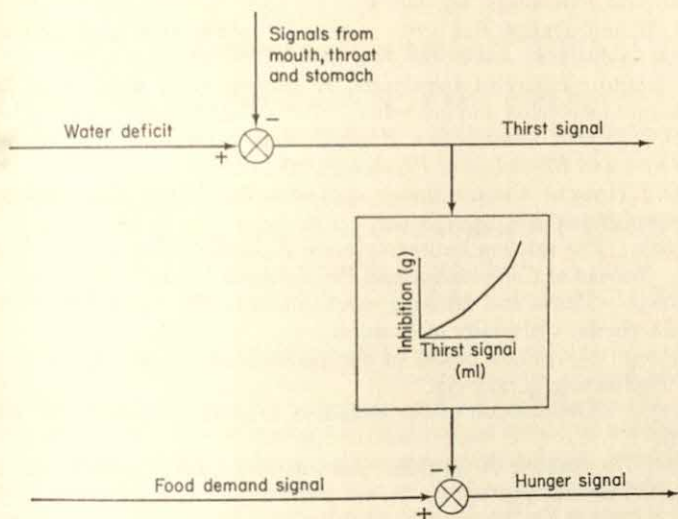
FIGURE 3.   A model of the effect of thirst on hunger, in terms of control theory.

from the amount that would otherwise have been eaten.   The non-linear inhibitory function represents an operation performed on the thirst signal, to reduce intake more severely at higher levels of water deficit or osmotic load.

In Experiment II, after 24 hr of food deprivation, disinhibition was at the rate of 0·2 g per ml, and after 48 hr 0·37 per ml of water drunk.   The difference between these values (which was significant at the 1% level) corresponds to the non-linearity discussed above.   However the values lie on either side of the mean (i.e. $\frac{1}{3}$ g food inhibited by 1 ml of water deficit), whereas if this approximate value found earlier corresponded to the slope of the lower part of the non-linear function, both the values found in Experiment II might have been expected to be somewhat higher with the value for 24-hr deprivation around 0·3.   This discrepancy is somewhat puzzling though the values are within the range of sample estimates of the disinhibitory factor made previously which have varied between 0·19 and 0·40.   Though this variance may seem large, it probably conceals factors (such as stomach contents) which we are not yet able to assess quantitatively.

References

GRIMSLEY, D. L. (1965).   Effects of water deprivation and injections of hypertonic saline on the activity of rats.   Psychological Reports, 16, 1081–5.
HSAIO, S. (1970).   Reciprocal and additive effects of hyperoncotic and hypertonic treatments on feeding and drinking in rats.   Psychonomic Science, 19, 303–4.
HSAIO, S. and KENDRICK, D. (1970).   Feeding-drinking interaction: Efficacy of saline solutions for ingestion of dry food in rats.   Acta Psychologica Taiwanica, 12, 62–7.
HSAIO, S. and LANGENES, D. J. (1971).   Liquid intake, initiation and amount of eating as determined by osmolality of drinking liquids.   Physiology and Behaviour, 7, 233–7.

HSIAO, S. and TRANKINA, F. (1969). Thirst-hunger interaction: I. Effects of body-fluid restoration on food and water intake in water-deprived rats. *Journal of Comparative and Physiological Psychology*, **69**, 448–53.

KAKOLEWSKI, J. W. and DEAUX, E. (1970). Initiation of eating as a function of ingestion of hypo-osmotic solutions. *American Journal of Physiology*, **218**, 590–5.

KUTSCHER, C. (1966). Effect of hypertonic saline injections and water deprivation on drinking, serum osmolality and gut water. *Physiology and Behaviour*, **1**, 259–68.

MCFARLAND, D. J. (1964). Interaction of hunger and thirst in the barbary dove. *Journal of Comparative and Physiological Psychology*, **58**, 174–9.

MCFARLAND, D. J. (1965). Control theory applied to the control of drinking in the barbary dove. *Animal Behaviour*, **13**, 478–92.

NOVIN, D. (1962). The relation between electrical conductivity of brain tissue and thirst in the rat. *Journal of Comparative and Physiological Psychology*, **55**, 145–54.

OATLEY, K. (1965). Thirst and drinking mechanisms in the regulation of water intake in rats. Ph.D. thesis, University of London.

OATLEY, K. (1967). A control model of the physiological basis of thirst. *Medical and Biological Engineering*, **5**, 225–37.

OATLEY, K. (1971). Dissociation of the circadian drinking pattern from eating. *Nature*, **229**, 494–6.

OATLEY, K. and DICKINSON, A. (1970). Air drinking and the measurement of thirst. *Animal Behaviour*, **18**, 259–65.

OATLEY, K. and TOATES, F. M. (1971). Inhibition of eating by osmotic thirst and dis-inhibition by drinking. Paper presented to 4th International Conference on the regulation of food and water intake.

OATLEY, K. and TONGE, D. A. (1969). The effect of hunger on water intake in rats. *Quarterly Journal of Experimental Psychology*. **21**, 162–71.

STELLAR, E. and HILL, J. (1952). The rat's rate of drinking as a function of water depriva-tion. *Journal of Comparative and Physiological Psychology*, **45**, 96–102.

TOATES, F. M. and OATLEY, K. (1970). Computer simulation of thirst and water balance. *Medical and Biological Engineering*, **8**, 71–87.

VALENSTEIN, E. S., COX, V. C. and KAKOLEWSKI, J. W. (1970). Re-examination of the role of the hypothalamus in motivation. *Psychological Review*, **77**, 16–31.

WAYNER, M. J. and PETRAITIS, J. (1967). Effects of water deprivation and salt arousal of drinking on wheel turning activity of rats. *Physiology and Behaviour*, **2**, 273–5.

ON THE DIVISION OF ATTENTION: A DISPROOF OF THE SINGLE CHANNEL HYPOTHESIS

D. ALAN ALLPORT, BARBARA ANTONIS AND PATRICIA REYNOLDS

Reading University

In dichotic listening, subjects are apparently unable to attend simultaneously to two concurrent, auditory speech messages. However, in two experiments reported here, it is shown that people can attend to and repeat back continuous speech at the same time as taking in complex, unrelated visual scenes, or even while sight-reading piano music. In both cases performance with divided attention was very good, and in the case of sight-reading was as good as with undivided attention. There was little or no effect of the dual task on the accuracy of speech shadowing. These results are incompatible with the hypothesis that human attention is limited by the capacity of a general–purpose central processor in the nervous system. An alternative, "multi-channel", hypothesis is outlined.

Introduction

Over the past decade, one hypothesis on the limits of human attention and performance has achieved something of the status of a central dogma. This is the hypothesis that the brain, in its moment-to-moment decision-taking regarding both immediate and potential future response, acts as a single communication channel of limited capacity (Broadbent, 1958; Craik, 1948; Welford, 1952). According to this view, performance of two or more concurrent tasks can only be maintained by the rapid alternation of attention (i.e. by time-sharing on the access to a general-purpose central channel) between the requirements of the different tasks. Where one of the tasks is such as to demand continuous occupation of the single channel, information relating to other tasks is necessarily excluded. Their performance fails. Alternatively, if information relating to a second task is admitted, performance of the first will be disrupted.

There have been numerous recent reaffirmations of this so-called single channel hypothesis (Broadbent, 1971; Deutsch and Deutsch, 1963; Kahneman, 1970; Kristofferson, 1967; Lindsay, 1970; Moray, 1967, 1969; Morton and Broadbent, 1967; Neisser, 1967; Norman, 1968; Smith, 1969; Swets and Kristofferson, 1970; Welford, 1968), although coupled with widely differing speculations regarding the origin of the supposed central bottleneck. Controversy in this area has been confined largely to two issues: (1) the level at which selection between competing messages or tasks becomes necessary (i.e. on the amount of pre-processing prior to access to the single channel) and (2) the adequacy with which competing or irrelevant messages are "rejected". Relating to (1), it has been widely accepted that, on the input side, different "analysers" can operate in parallel on different aspects

of the sensory data. (Thus, in the Conclusions to her lucid and oft-quoted review, Treisman (1969) writes: "These findings suggest that division of attention between two or more inputs and between two or more targets is difficult or impossible when no time is allowed for alternating attention or serial analysis . . . while division of attention between analysers is relatively efficient . . .", p. 296.) It has also been suggested (e.g. Shaffer, 1971) that, in a continuous task, simultaneous or over-lapping processing of input and output can occur. On the other hand the basic postulate of a general-purpose central processor of limited capacity at *some* stage ("categorization", "translation"?) between input and output as the basis of human attention limitations does not appear to have been seriously challenged, at least in recent years.

(The single channel hypothesis has a weak and a strong version. The strong version asserts that no two independent sensory messages can be handled concurrently. The weak version asserts simply that there is a limited processing capacity to a single—i.e. general-purpose—"central channel"; so long as this is not exceeded, more than one signal may be dealt with at the same time. The experiments described here are directed at the weaker version, any refutation of which is *a fortiori* a rejection of the strong version also.)

The most striking evidence put forward in support of the single channel hypothesis concerns the inability of human subjects to take in more than one verbal message at a time: the "cocktail party problem" (Cherry, 1953). Commonly, experiments on this problem have included the requirement that the subject also continuously repeat back, or "shadow", one of the messages as it is presented, to ensure that he is objectively attending to at least one of them. In this situation the subject is unable to report or recognize afterwards any of the contents of the message he had not been shadowing, nor even what language it was spoken in, at least for any part of the non-shadowed message presented more than a few seconds prior to an interruption of shadowing, (Glucksberg and Cowen, 1970; Moray, 1959; Norman, 1969; Mowbray, 1964). Analogous results have been obtained when parallel lines of printed text are presented for reading (Neisser, 1969), and when different prose passages are presented simultaneously to eye and ear (Mowbray, 1953). In the context of the single channel hypothesis, the interpretation of these results which presented itself was therefore that shadowing one continuous language input satisfied the theoretical condition of full, continuous occupation of the limited capacity single channel.

One thing that stands out about these and other experimental paradigms (e.g. Moray, 1969) which have been employed in support of the single channel hypothesis is that they call for simultaneous attention to two closely similar, or even identical, tasks. That is, either the sensory inputs concerned, or more importantly the functions to be computed on them, appear similar in kind. It is possible, therefore, that the difficulty in such cases derives not from exceeding the limited capacity of a single, general-purpose central processor, but more simply from the difficulty of keeping separate (i.e. of not confusing or confounding) two closely similar but unrelated messages. In either case the subject's only workable strategy may be to filter or reject one of them. On the other hand where the messages, or the tasks to be performed, are highly dissimilar, we suggest (Allport, 1971b) this difficulty

should not arise, and both tasks should be able to be performed simultaneously. We therefore looked for situations involving highly dissimilar tasks, but of comparable information content, which could provide a critical test of the single channel hypothesis, and which at the same time might throw some light on the conditions in which division of attention is or is not possible.

Two experiments are described. In both, auditory speech shadowing was one of the concurrent tasks, since it is generally represented, on the evidence just reviewed, as the paradigm of tasks requiring complete and uninterrupted occupation of the hypothetical single channel. It also has the advantage that details of performance at this task are relatively well documented.

Experiment I

In the first experiment a monaural, auditory shadowing task was combined with simultaneous presentation of either verbal or non-verbal stimuli to be memorized (words or pictures). The words were presented, in different conditions, either auditorily or visually.

Method

Materials

The texts for auditory shadowing were passages from George Orwell's "Selected Essays" recorded in a female voice on a Ferrograph twin-track tape-recorder, and delivered to the subject's right ear by headphones. Each passage was about 180 words long and lasted 1 min. Another tape recorder monitored the subject's shadowing performance. The competing memory stimuli were of three kinds: (1) A list of 15 words, recorded by a male voice at the rate of one every 3 sec, delivered to the subject's left ear by well-padded headphones. This condition was modelled on experiments previously described by Moray (1959), and others. (2) Words presented visually, in upper case type, also at the rate of one every 3 sec. (3) A series of 15 coloured photographs, displayed at the same rate. The words for auditory and visual presentation were concrete nouns of 2–3 syllables, having a frequency of one per million in the Thorndike–Lorge word list (Thorndike and Lorge, 1944). The pictures were complex scenes selected from coloured magazines, to have as varied as possible subject matter; they had not previously been seen by any of the subjects. (We selected each set of memory items from a pilot experiment, using other subjects and without the concurrent shadowing task, to obtain equal levels of difficulty—i.e. equal probability of subsequent correct recognition—both between different sets of the same type and, in particular, between the different types of memory item. The same set of 15 memory items was of course presented once only to each experimental subject.) A Kodak Carousel slide projector exposed the visual items for 1·7 sec each, with 1·3 sec between exposures, on a white screen 45 cm by 60 cm. The words occupied about one third of the screen in horizontal extent, or approximately 7 degrees of visual angle (at a viewing distance of 160 cm).

Procedure

We tested how much our subjects had taken in of the memory items by a forced-choice recognition procedure. After each 1-min trial, there was an interval of 1 min and then the 15 test items just presented were again shown with 15 new items drawn from the same population; the 30 items were presented singly in random order. The subjects responded to each item either "Yes", if they believed it to be one of the original test items, or "No", if not. They were given no indication of their accuracy. Recognition memory was tested in this way for both divided and undivided attention, i.e. both with, and without, concurrent

auditory shadowing during presentation of the memory items. In the dual task condition it was stressed that the shadowing task was to be treated as primary.

On the first day the subjects practised auditory shadowing, without competing inputs, to a criterion of one 1-min passage repeated without errors (words omitted or confused). Experimental testing was carried out on the following day in two 45-min sessions for each subject. The order in which the different conditions were encountered was counterbalanced across subjects, and between the first and second experimental session. In the first experimental session a trial under undivided attention was immediately followed by one trial of the corresponding dual task condition, and so on, through the three different types of memory item. In the second session the order of alternation was reversed. In the dual tasks, presentation of the memory items began 7·5 sec after the start of the shadowing text and terminated 7·5 sec before the text came to an end.

Subjects

The subjects were six female undergraduates. They were told that they were participating in a study of dual task performance, but were otherwise not informed about the nature of the experiment. They were encouraged to perform as well as they could, and appeared to be highly motivated and to enjoy testing themselves in the experiment.

Results

Figure 1(a) shows the results. Recognition memory for auditory words presented during shadowing approached chance level (50% errors). This is the result obtained by other experimenters (Glucksberg and Cowen, 1970; Moray, 1959;
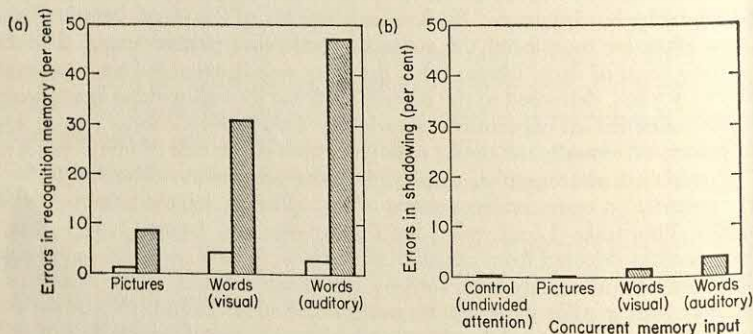
FIGURE 1. Accuracy of performance (Experiment I) is shown as percentage error: (a) for recognition memory of pictures and words presented both with and without concurrent auditory shadowing; (b) for the shadowing task, both in undivided attention (control) and concurrently with presentation of each of the three modes of memory items. □ Undivided attention. ▨ With concurrent shadowing.

Norman, 1969; Mowbray, 1964). Subsequent recognition of *visually* presented words was also greatly affected by shadowing during presentation; all subjects on every trial made more errors in recognizing the visual words under divided attention than under undivided attention. (Questioned about their performance after the experiment, the subjects claimed to have recognized the visual words, at least in part, on the basis of their *visual* characteristics: number of letters, presence of repeated letters, small marks on the slides. Had they been forced to respond on the basis of semantic or phonological rather than visual properties of the words,

memory for visually presented words with concurrent shadowing might perhaps have been as poor as for the words presented auditorily).   Recognition of the pictures was much less affected by simultaneous auditory shadowing; one subject showed *no* loss of accuracy, and three made only one or at most two more errors under divided than under undivided attention.   The mean deterioration in picture recognition under divided attention was, however, statistically significant ($t = 2\cdot71$, $P < 0\cdot05$).   On the other hand, the difference between the three stimulus modes in the extent of deterioration was highly significant ($F = 17\cdot5$, $df = 2,10$, $P < 0\cdot001$), and between each mode individually ($P < 0\cdot01$ in each case, by a Newman–Keuls multiple comparison test: auditory words > visual words > pictures) There were no consistent differences in performance between sessions 1 and 2. ($F < 1\cdot0$, $df = 1,5$).

These differences in the extent of deterioration cannot be attributed simply to differences in the intrinsic difficulty of the particular memory tasks, as evidenced under *undivided* attention.   Visually presented words were *less* accurate in undivided attention than auditorily presented words, but nevertheless suffered significantly less deterioration in accuracy under divided attention.   Neither can they be attributed to a trade-off with the accuracy of concurrent shadowing (Fig. 1(b)). While the type of memory input has a significant effect on shadowing accuracy ($F = 7\cdot25$, $df = 2,10$, $P < 0\cdot05$) the incidence of shadowing errors is positively related to the extent of deterioration in the concurrent memory task.

Discussion

These results are not easily reconciled with the single channel hypothesis. According to most versions of this hypothesis (e.g. Broadbent, 1971; Neisser, 1967; Welford, 1968), inputs to long-term memory must pass through the general-purpose "single channel".   If this is already fully occupied by the shadowing task, then merely altering the modality or other characteristics of a secondary input could not affect the latter's chances of gaining entry to long-term memory (provided, of course, for the weak version of the hypothesis, that the information content of the secondary input is not also drastically reduced; this latter does not appear a very promising way out for the hypothesis in the present case).   It cannot be argued in defence of the single channel hypothesis that the shadowing task, even though at a very fast rate, might not have been such as to fully occupy the capacity of the hypothetical single channel, since in that case the hypothesis is unable to account for the almost complete failure to remember the auditory word list (and to a lesser extent the visual word list) under the same conditions of concurrent shadowing during presentation.   Neither does it appear plausible to argue that very rapid switching or time-sharing on the single channel may be possible as between the processing of pictures and the shadowing task, but not between discrete auditory words and shadowing.   If the processing of visual input is intrinsically very much faster than that of auditory inputs, then how is the relatively large deterioration in the case of visually presented words, under divided attention, to be explained?

Experiment II

The second experiment investigated the simultaneous performance of continuous auditory prose shadowing and playing piano music from a score.

Method

Materials

The shadowing task was similar to that in Experiment I, except that recorded prose passages, of approximately 1 min duration, were presented at 150 words per min. Passages were of two levels of difficulty: "easy" texts were selected from an anthology of humorous narrative prose, containing predominantly high-frequency words; "difficult" texts were taken from a textbook of early Norse history, containing a high proportion of low-frequency words. The music for sight-reading consisted of examination pieces, Grades II and IV, of the Associated Board of the Royal Schools of Music (1970 and 1971), none of which had previously been seen by the subjects.

The prose for shadowing was delivered binaurally by headphones, and the subjects' performance on the shadowing and piano playing tasks was recorded on separate tracks of a tape recorder. In the dual task conditions, the prose passage commenced 10 sec after the subject had started sight-reading. The instructions were to continue as evenly and accurately as possible on both tasks, and not to try to correct errors. Neither task was presented as primary. The subjects were also warned that, following some of the trials, they would be asked questions about the content of the prose passages they had just shadowed.

Procedure

Subjects received ten 1-min passages of practice at auditory shadowing with easy narrative prose. All subjects reached the criterion of two successive trials without omissions in this period. They then received two trials of practice at sight-reading, one each at Grades II and IV, and finally seven trials at the combined task of shadowing and piano sight-reading simultaneously. On the day following practice, the subjects served in two experimental sessions of sight-reading both with and without concurrent auditory shadowing. In session 1 the sight-reading alone was performed first (i.e. the subject ignored the simultaneously presented prose passage) and the dual task second; in session 2 this sequence was reversed. Each subject performed two trials of the dual task at each possible combination of "easy" or "difficult" prose and Grade II or IV of piano music. The subjects were five third-year undergraduate Music students.

Results

The accuracy of prose shadowing with concurrent sight-reading is shown in Figure 2(a). In session 1 the overall frequency of errors is affected both by the difficulty of the prose to be shadowed and by the Grade of concurrent sight-reading required. By session 2, however, the latter effect has entirely disappeared. Notice that only shadowing of the "easy" prose is to be directly compared with the criterion performance (zero omissions) in undivided attention, since this was obtained for "easy" texts only. The majority of errors were syllable or word repetitions (stammering), and mispronunciations. The few omission errors with "easy" text under divided attention were due to two subjects. The remainder shadowed without any omissions. None of the error rates with "easy" prose were significantly different from the criterion performance under undivided attention, except in session 1 when combined with Grade IV sight-reading ($t = 4.48$, $P < 0.05$). As a further check on the possibility of rapid alternation of attention between the shadowing and sight-reading tasks, and since timing is a critical feature of the latter, we also examined the fidelity of shadowing in the timing of inter-word intervals relative to the recording they shadowed. One trial was taken from each subject's records under divided attention, and the time intervals between successive
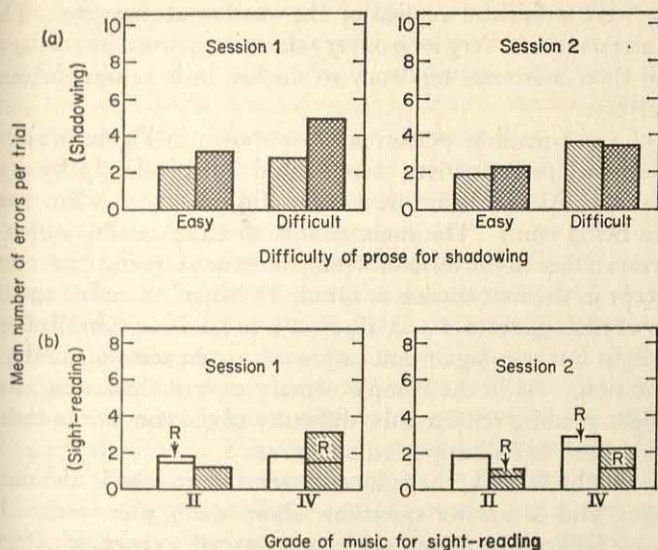
FIGURE 2. (a) Accuracy of auditory shadowing (Experiment II) is shown in terms of the mean number of errors per one-minute trial, for both "easy" and "difficult" prose, and with two grades of difficulty of concurrent sight-reading. ▨ Grade II of concurrent sight-reading ▨ Grade IV of concurrent sight-reading. (b) Accuracy of sight-reading piano music of two grades of difficulty is shown in terms of the mean number of errors per trial (max > 300) both with and without concurrent auditory shadowing. Errors marked "R" above the horizontal bar on each column refer to rhythm or timing errors in performance; those below the bar are wrong notes. ☐ Undivided attention; ▨ With concurrent auditory shadowing.

word onsets both for the original tape on that particular trial and for the subject's shadowing output were obtained. This was approximated by replaying the recordings at the slowest tape speed and keying word onset times manually on a pen recorder. The results, summarized in Figure 3, showed that the time relations
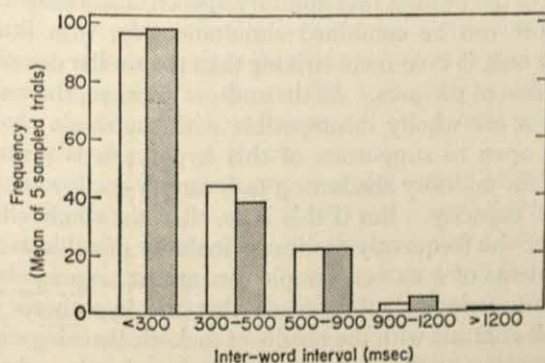
FIGURE 3. Inter-word intervals (onset—onset) in both the original tape-recorded text for shadowing (the "model") and in the subjects' repetition of it while simultaneously sight-reading on the piano. ☐ Model. ▨ Subject.

in shadowing were a faithful replica of the shadowed message. There was no indication of an increase in very long or very short inter-word intervals, which would be expected if there were any tendency to shadow in intermittent bursts of rapid speech and silence.

Accuracy of sight-reading performance is shown in Figure 2(b). (Each subject's tape-recorded performances were scored independently by a member of Reading University Music department. He did not know, when scoring, which condition was being run.) The main feature of these results is that there is no increase in errors either in the form of wrong notes or of timing errors under divided attention, except in the first session at Grade IV where there is a small increase in the frequency of timing errors ($t = 8\cdot18, P < 0\cdot01$). Indeed in all other conditions there was a slight but non-significant improvement in accuracy under the divided attention condition. As in the complementary case of shadowing, the frequency of errors in sight-reading reflected the difficulty of the concurrent task during session 1, but this effect had disappeared by session 2.

During the combined task, the subjects were able to take in the meaning of the prose messages, and to answer questions about them afterwards. Immediately after shadowing Difficult prose our subjects answered on average 61% of questions correctly when the simultaneous sight-reading had been at Grade II, and 56% at Grade IV. The difference was not significant. The same questions presented to control subjects, who had not heard the prose passages concerned and were asked to guess the answers, were answered 12% correctly. Subjects shadowing the same passages with undivided attention gave 57% correct answers. The difference between memory for the shadowed text in divided and undivided attention was not statistically significant. The most remarkable feature of these results, however, was the wide variation between the different experimental subjects. One, the most competent pianist in the group, scored 81% correct. The least proficient sight-reader among our subjects only answered 14% of the same questions correctly.

Discussion

In many respects the finding that auditory speech shadowing and playing piano music from a score can be combined simultaneously, with little or no loss of efficiency in either task, is even more striking than the similar demonstration regarding the memorization of pictures. In the authors' opinion, the results of the sight-reading experiment are wholly incompatible with the single channel hypothesis. The only option open to supporters of this hypothesis is presumably to argue *ex post facto* that the auditory shadowing task cannot occupy all, or even most, of the single channel capacity. But if this is so, then the single channel hypothesis fails to account for the frequently replicated inability of subjects while shadowing to take in the contents of even very simple concurrent language inputs, which was after all one of the principal phenomena that the hypothesis was designed to explain. It is this contrast with the results of dichotic listening experiments which is of primary importance; the question of whether the time sharing is complete or not is secondary to the main theoretical argument. The word rate and relative unpredictability of the prose used for shadowing in the present experiment are at least as great as that in other experiments in which the, now classical, failure to

divide attention between competing speech inputs was found, and the accuracy of shadowing by our subjects was at least as good as in these experiments. If there is, in the functional organization of the brain, anything analogous to the general-purpose central processor of the von Neumann computer, we are forced to conclude that it is not involved in the performance of these tasks.

A more appropriate model would be that of a number of independent, special-purpose computers (processors and stores) operating in parallel and, at least in some cases, capable of accepting only one message or "chunk" of information for processing at a time. In general, we suggest, any complex task will depend on the operation of a number of independent, specialized processors, many of which may be common to other tasks. To the extent to which the same processors are involved in any two particular tasks, truly simultaneous performance of these two tasks will be impossible. On the other hand, the same tasks paired respectively with another task requiring none of the same basic processors can in principle be performed in parallel with the latter without mutual interference. We will refer to this general class of model, for convenience, as the "multi-channel" hypothesis.

Clearly, if we are correct, the mistake made in postulating the generalized single channel hypothesis was to assume that the information-processing bottlenecks, for which evidence could be found in a number of different experimental situations, were manifestations in each case of the same, "central" bottleneck. Seen in this light, there is nothing particularly surprising, or even novel, about the multi-channel hypothesis. It agrees well with what is known of the gross functional anatomy of the brain (Luria, 1970), and, so far as it goes, with single unit physiology. It is also in general agreement with the evidence of simultaneous processing of stimuli varying along different attributes or "dimensions" within the same sense mode (Allport, 1971a; Hawkins, 1969; Marcell, 1970; and others), and with the elegant demonstrations by Brooks of interference between reading and visualizing the referent of a verbal description (Brooks, 1967, 1970). An alternative and simpler explanation of the results of Experiment I might be given in terms of functional independence between the two cerebral hemispheres, with language processing in the left or "dominant" hemisphere, picture processing in the right. This interpretation could only with difficulty be extended to the second experiment, at least in regard to the control of right-hand fingering in piano playing. It is altogether implausible if applied to the results, just cited, showing full parity of performance in the simultaneous processing of, for example, colour and form information from tachistoscopic exposures.

We do not wish to deny that the brain may, in certain circumstances, exhibit "single channel" operation as a whole. This may occur when someone concentrates on a particular task: most, or all, of the specialized processors are being held "on call" to the same message source whether they are in fact being used or not. We deny merely that this organization is obligatory. In a pilot study for Experiment I we observed in ourselves, as subjects, a strong initial disposition to look away or close our eyes while shadowing, in order not to be distracted by the intruding picture series. It came as a surprise to discover, on the first attempt, that division of attention between these two inputs was in fact quite easy. The more unpredictable (to the subject) the demands of a particular task, or the greater the

penalties for errors in its performance, the less the probability that multi-channel or distributed processing of other tasks, even though possible, will be permitted. For this reason, *failure* to obtain efficient performance of two concurrent tasks is not in itself unambiguous evidence that the functional "capacity" for both is lacking; a change of payoff may produce a quite different result (e.g. Kahneman, 1970). If, as seems plausible, one component in the acquisition of a skill is an improvement in predicting, or discriminating, the demands which are likely to be made by the task, then, with increasing skill, functions not in fact necessary to the performance should no longer be held on call, and so pre-empted from other tasks: with increasing skill, less "concentration" is needed. Similar predictions concerning the effects of practice on the division of attention have sometimes been claimed for the single channel hypothesis, and any hypothesis on the nature of attention must of course be able to deal with these, often very striking, effects. We include the foregoing speculative remarks merely to show that the single channel hypothesis has no special claim in this respect.

References

ALLPORT, D. A. (1971a). Parallel encoding within and between elementary stimulus dimensions. *Perception and Psychophysics*, **10**, 104–8.

ALLPORT, D. A. (1971b). A new hypothesis on the nature of attention. Paper presented at meeting of Experimental Psychology Society, Reading, March, 1971.

BROADBENT, D. E. (1958). *Perception and Communication*. London: Pergamon Press.

BROADBENT, D. E. (1971). *Decision and Stress*. London: Academic Press.

BROOKS, L. R. (1967). The suppression of visualization by reading. *Quarterly Journal of Experimental Psychology*, **19**, 289–99.

BROOKS, L. R. (1970). An extension of the conflict between visualization and reading. *Quarterly Journal of Experimental Psychology*, **22**, 91–6.

CHERRY, C. (1953). Some experiments on the reception of speech with one and with two ears. *Journal of the Acoustical Society of America*, **25**, 975–9.

CRAIK, K. J. W. (1948). Theory of the human operator in control systems: II. Man as an element in a control system. *British Journal of Psychology*, **38**, 142–8.

DEUTSCH, J. A. and DEUTSCH, D. (1963). Attention: some theoretical considerations. *Psychological Review*, **70**, 80–90.

GLUCKSBERG, S. and COWEN, G. N. (1970). Memory for nonattended auditory material. *Cognitive Psychology*, **1**, 149–56.

HAWKINS, H. L. (1969). Parallel processing in complex visual discrimination. *Perception and Psychophysics*, **5**, 56–64.

KAHNEMAN, D. (1970). Remarks on attention control. *Acta Psychologica*, **33**, 118–31.

KRISTOFFERSON, A. B. (1967). Successiveness discrimination as a two-state quantal process. *Science*, **158**, 1337–40.

LINDSAY, P. H. (1970). Multichannel processing in perception. In MOSTOVSKY, D. I. (Ed.), *Attention: Contemporary Theory and Analysis*. New York: Appleton-Century-Crofts.

LURIA, A. R. (1970). The functional organization of the brain. *Scientific American*, **222**, 66–78.

MARCEL, A. J. (1970). Some constraints on sequential and parallel processing, and the limits of attention. *Acta Psychologica*, **33**, 77–92.

MORAY, N. (1959). Attention in dichotic listening: affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology*, **11**, 56–60.

MORAY, N. (1967). Where is capacity limited? A survey and a model. *Acta Psychologica*, **27**, 84–93.

MORAY, N. (1969). *Attention: Selective Processes in Vision and Hearing*. London: Hutchinson.

MORTON, J. and BROADBENT, D. E. (1967). Passive versus active recognition models, or is your homunculus really necessary? In WATHEN-DUNN, W. (Ed.), *Models for the Perception of Speech and Visual Form*. New York: M.I.T. Press.

MOWBRAY, G. H. (1953). Simultaneous vision and audition: the comprehension of prose passages with varying levels of difficulty. *Journal of Experimental Psychology*, **46**, 365–72.

MOWBRAY, G. H. (1964). Perception and retention of verbal information presented during auditory shadowing. *Journal of the Acoustical Society of America*, **36**, 1459–65.

NEISSER, U. (1967). *Cognitive Psychology*. New York: Appleton-Century-Crofts.

NEISSER, U. (1969). Proceedings of XIX International Congress of Psychology, London.

NORMAN, D. A. (1968). Toward a theory of memory and attention. *Psychological Review*, **75**, 522–36.

NORMAN, D. A. (1969). Memory while shadowing. *Quarterly Journal of Experimental Psychology*, **21**, 85–93.

SHAFFER, L. H. (1971). Attention in transcription skill. *Quarterly Journal of Experimental Psychology*, **23**, 107–12.

SMITH, M. C. (1969). The effect of varying information on the psychological refractory period. *Acta Psychologica*, **30**, 220–31.

SWETS, J. A. and KRISTOFFERSON, A. B. (1970). Attention. *Annual Review of Psychology*, **21**, 339–66.

THORNDIKE, E. L. and LORGE, I. (1944). *The Teacher's Word Book of 30,000 Words*. New York: Teachers College.

TREISMAN, A. M. (1969). Strategies and models of selective attention. *Psychological Review*, **76**, 282–99.

WELFORD, A. T. (1952). The "psychological refractory period" and the timing of high-speed performance—a review and a theory. *British Journal of Psychology*, **43**, 2–19.

WELFORD, A. T. (1968). *Fundamentals of Skill*. London: Methuen.

BOOK REVIEWS

DIXON, N. F.   *Subliminal Perception: the Nature of a Controversy*.   London: McGraw-Hill.
1971. Pp. 363. £4.00.

A sane attitude towards the question of subliminal perception appears now to be supplant-
ing the extravagant positive and negative attitudes of the past 20 years, and Dr Dixon's
absorbing book has thus arrived at an opportune moment.   The book begins with a brief
historical survey, and then takes up the question of the definition of subliminal perception.
This definition is formulated in terms of three necessary criteria.   Two of these are straight-
forward.   The third, however, seems an odd one; subliminal perception is said necessarily
to involve "the occurrence of contingent responses, without reported awareness of the
stimulus, that differ qualitatively from those elicited by the same stimulus when presented
above the awareness threshold" (p. 18).   Now, while it may turn out to be the case that this
is how subliminal perception works (and much evidence is presented in favour of this view
later in the book), surely the property in question is accidental rather than essential.   There
seems no reason to exclude, as instances of subliminal perception, responses to a subliminal
stimulus which are qualitatively similar to the responses made to the stimulus when it is
supraliminal.

Some telling points are made in this part of the book.   The naïveté of the early attempts
to "solve the problem" of perceptual defence by "converging operations" is pointed out,
as is the foolishness of the early logical objections to the notion of subliminal perception,
which argued that it could not occur because it is by definition self-contradictory.

Subsequent chapters are devoted to closely argued and extremely detailed reviews of a
variety of topics related to subliminal perception: the effects of subliminal anchors and
subliminal cues, the Poetzl phenomenon, the problem of differing drive states, perceptual
defence, possible physiological bases for subliminal perception, and selective attention.

Dr Dixon devotes considerable time to discussing the views of Neisser on subliminal
perception and in particular the relationship between iconic memory and subliminal percep-
tion.   His discussion of these views, however, appears to me to be vitiated by two serious
misunderstandings of Neisser's position.   First, he claims (p. 56) that Neisser (1967) holds
that icons "are consciously held visual images".   This statement is misleading because it
neglects Neisser's distinction between preattentive and focal processes.   One is aware that
an icon is present, but not of the nature of its contents.   The latter must await focal process-
ing.   One is only aware of iconic memory in the sense that crude information about the
properties of its contents (colour, spatial location, etc.) are available.   To be aware of the
identity of items in iconic memory one must first translate them to a different form of storage.

Nor is it correct to claim, as Dr Dixon does, that Neisser believes that iconic memory is
an essential link between visual input and identification.   Iconic memory is only needed
when the visual stimulus is terminated before it has been adequately processed; in this case,
processing of the stimulus can continue using iconic memory as a source of data.   If,
however, the visual stimulus is present for long enough, its subsequent iconic memory is
irrelevant, since processing will have been completed before stimulus offset.

It is of course the case that Neisser expresses doubts about the notion of subliminal
perception, but this is for other reasons; there is no inconsistency between Neisser's con-
ception of iconic memory and the existence of subliminal perception.   The processes of
analysis-by-synthesis and verbal encoding by which Neisser considers items in iconic
memory are identified are not obligatorily conscious processes.

The book concludes with an explicit statement of the author's view of phenomenal
experience as a process which occurs "in parallel with (the processes) mediating between
stimuli and overt responses" (p. 308).   It seems to me that this view is inexactly stated, and,
in particular, that the term "in parallel" is incorrectly used.   As is made clear by Figure

1.1b, the author's view is that phenomenal experience (i.e. awareness, consciousness), if it occurs in response to a stimulus, is a *consequence* of those internal processes which link stimulus to response. The crux of this view is that awareness is neither a necessary component, nor a necessary product, of the linkage between stimulation and behaviour. Thus awareness is not considered to be *in parallel* with these linking processes; both awareness and behaviour are *consequences* of these linking processes, and either can occur without the other, although behaviour can be influenced by awareness.

Since, on this view, reaction to stimulation does not require awareness, one is tempted to ask what, if anything, awareness is for. The question is treated only in an indirect way by asserting that a subliminal perception is by definition qualitatively different from a supraliminal perception; it might be argued that, for reaction to stimulation to be appropriate or rational, awareness is required. However, as argued above, such a definition seems unduly arbitrary.

Consideration of current work on visual information-processing benefits from an acquaintance with Dr Dixon's book. For example, Wheeler (1970) has shown that the exposure duration required for a word to reach awareness is less than that required for a non-word. To argue that this is because the word is meaningful would seem to invite the rejoinder that the meaning of a word can only be known *after* the word is identified, in which case it is hard to see how identification could be *assisted* by meaning. Perhaps this is why Wheeler attempted to explain the effect without involving meaning; but surely it is more satisfactory to argue, fortified by Dr Dixon's book, that the meaning of an item can be available prior to one's being aware of the item, in which case it is no longer difficult to see how the meaning of an item could contribute to one's becoming aware of it.

A similar point can be made in connection with the demonstration by Eriksen, Pollack and Montague (1970) that reaction time in a number-naming task is positively related to the number of syllables in the name of the number. These authors interpret this finding as evidence that inner speech precedes, and is a prerequisite for, identification; that such a view clearly implies some form of subception acquires piquancy from the fact that, as Dr Dixon notes, Eriksen was favourably disposed towards subception in 1954 but unfavourably disposed towards it in 1960.

The book, then, has its virtues and its defects. Among its virtues are its spirited defence of subliminal perception, and its eclectic and scholarly analyses of the literature. Its treatment of visual information-processing, on the other hand, is marred by an idiosyncratic interpretation of Neisser's views; and one would have liked more extensive discussions of the author's theoretical attitude towards subliminal perception. To say that subliminal perception exists is not to say anything about how it operates.

Page 49 contains what is either an exquisite misprint or an abominable pun.

References

ERIKSEN, C. W., POLLACK, M. D. and MONTAGUE, W. E. (1970). Implicit speech: a mechanism in perceptual encoding? *Journal of Experimental Psychology*, **84**, 502–7.

NEISSER, U. (1967). *Cognitive Psychology*. New York: Appleton-Century-Crofts.

WHEELER, D. D. (1970). Processes in word recognition. *Cognitive Psychology*, **1**, 59–85.

MAX COLTHEART

ARONSON, L. R., TOBACH, E., LEHRMAN, D. S. AND ROSENBLATT, J. S. (Eds). *Development and Evolution of Behavior*. San Francisco: W. H. Freeman and Company. 1970. Pp. 656 + xviii. £5.60.

This book contains 27 essays written in memory of T. C. Schneirla preceded by a brief biography of Schneirla, and a summary of his major contributions to important issues in animal behaviour. Much of his theoretical effort was devoted to the analysis of two distinctions, one, which he rejected, between innate and learned behaviour, and the other, which he propounded, between the processes underlying approach and withdrawal behaviour. Many of the essays provide strong support for his opposition to the concept of

innateness, although his contrast between approach and withdrawal has to date received rather less attention.

The book is divided into five parts, the last of which concerns human behaviour; the first four parts contain essays that cover a broad range of topics in the ethological field, and which are of considerable interest to all psychologists concerned with animals.

A recurring theme is the importance of the careful analysis of concepts intended for use in scientific argument. Lehrman's essay on the nature/nurture problem is largely concerned with establishing what is meant by many of the terms used by ethologists, and by Lorenz in particular. This chapter is a clear and convincing restatement of Lehrman's position, and is complemented by Gottlieb's essay, which reviews the theories and facts of the ontogeny of organized movement. Other essays that examine broad theoretical issues are those by Atz, on the establishment of behavioural homologies, by Vowles, on the analysis of the sequential organization of behaviour, and by Hinde and Stevenson, on the concepts of purpose and goal-directedness. All of these authors have made full use of this opportunity to develop their ideas in print.

Most of the remaining chapters in the first four parts of the book concern more specific pieces of behaviour; there are, for example, chapters on the physiological coding of releasing stimuli, on imprinting, on sexual and maternal behaviour in rats, on pillaging by stingless bees, and on animal navigation. Outstanding among these essays are chapters by Itô, Shaw, and Haskins. Itô's essay begins with a critical discussion of the concept of dominance-hierarchy, and broadens into an interesting and speculative account of the evolutionary and ecological bases of social organization in vertebrates. Shaw's essay, on schooling in fishes, follows a similar pattern, initially discussing the difficulties of distinguishing between schools and disoriented aggregations of fishes, before going on to speculations on the origins and value of schooling. Haskins' essay, on the origins of the nuptial flight and subsequent isolation of females in ants, beautifully illustrates the value of using the behaviour of living but primitive species to provide evidence for the way in which the behaviour seen in the most developed species may have evolved.

There are only six chapters on Human Behaviour, and one of these concerns the development of modern techniques of exhibiting animals in zoos. Inevitably, these essays appear a rather arbitrary appendage to the impressively broad ethological sections. Perhaps the most entertaining essay in this section is Maier's, on group problem-solving, a situation in which two heads ought, it seems, to be better than one, but frequently turn out to be rather worse. Other contributions discuss the classification of temperament in children, cognitive patterning in blind children, intergroup aggression, and psychoanalytic views of the nature/nurture problem.

This book, then, is a collection of highly readable essays that provide a broad coverage of theoretical topics (with the exception of the drive concept) that are current in ethology. It is well produced, with three indices (author, subject, and species), and is, for its length and quality, reasonably priced.

<div align="right">E. M. MACPHAIL</div>

GRUSSER, O. J. AND KLINKE, R. (Eds). *Pattern Recognition in Biological and Technical Systems*. Berlin: Springer-Verlag. 1968. Pp. 413.

This book contains the Proceedings of the Fourth Congress of the Deutsche Gesellschaft für Kybernetik. About three quarters of it are devoted to visual pattern recognition and the remainder to auditory recognition. It includes work on recognition by machine as well as papers on both psychological and neurophysiological aspects of pattern recognition. Almost none of the material is new; the topics covered seem to be a very haphazard collection, and most of the individual papers are themselves unsystematic reviews of an area and tend to deal in generalities.

The approach taken by almost all the authors is curiously dated and does not get beyond recognition by the detection of two-dimensional features or by the partitioning of an n-dimensional space. There is no mention of recent work in artificial intelligence on picture

processing and structural descriptions as undertaken by such men as Clowes, Minsky and Narasimhan. With the exception of articles by Julesz (on stereopsis), Richards (on size constancy scaling) and Harmon (on facial recognition) there is nothing on human pattern recognition. There are a large number of articles on receptive fields in the visual system, most of which repeat one another and throw little new light on the subject.

The book might have been worth publishing ten years ago; today it merely illustrates that the ideas current in cybernetics in the forties and fifties have no relevance to the problem of pattern recognition.

N. S. SUTHERLAND

Publications Received

BARBER, T., DICARA, L. V., KAMIYA, J., MILLER, N. E., SHAPIRO, D. AND STOYVA, J. (Eds). *Biofeedback and Self-control*. Chicago: Aldine-Atherton. 1971. Pp. 806. $17.50.

CARSWELL, E. A. AND ROMMETVEIT, R. (Eds). *Social Contexts of Messages*. London: Academic Press. 1971. Pp. 163. £2.75.

CHOMSKY, C. *The Acquisition of Syntax in Children from 5 to 10*. New York: M.I.T. Press. 1969. Pp. 124. £2.80.

CORCORAN, D. W. J. *Pattern Recognition*. Penguin Books. 1971. Pp. 223. £0·50.

EIBL-EIBESFELDT, I. *Love and Hate*. London: Methuen. 1971. Pp. 263. £2.50.

EYSENCK, H. J., ARNOLD, W. AND MEILI, R. (Eds). *Encyclopedia of Psychology*. London: Search Press. 1972. Pp. 396. £8.00.

FURCHTGOTT, E. (Ed). *Pharmacological and Biophysical Agents and Behavior*. New York: Academic Press. 1971. Pp. 401. $14.50.

HAMMER, A. G. *Elementary Matrix Algebra for Psychologists and Social Scientists*. Melbourne: Pergamon Press Australia. 1971. Pp. 212. £2.75.

HASSENSTEIN, B. *Information and Control in the Living Organism*. London: Chapman & Hall. 1971. Pp. 159. £1.30.

HECAEN, H. *Neuropsychologie de la perception visuelle*. Paris: Masson & Cie. 1972. Pp. 320. 110 F.

LADER, M. AND MARKS, I. *Clinical Anxiety*. London: William Heinemann. 1972. Pp. 202. £2.50.

LIEBERMAN, B. (Ed.). *Social Choice*. London: Gordon Breach. 1971. Pp. 425. £8.10.

MANIS, M. *An Introduction to Cognitive Psychology*. Wadsworth Publishing Co. Inc. 1971. Pp. 318. £4.75.

MCFARLAND, D. J. *Feedback Mechanisms in Animal Behaviour*. London: Academic Press. 1971. Pp. 279. £3.80.

MENYUK, P. *Sentences Children Use*. New York: M.I.T. Press. 1969. Pp. 162. £3.50.

MYERS, R. D. (Ed.). *Methods in Psychobiology*. London: Academic Press. 1971. Pp. 356. £6.00.

STROH, C. M. *Vigilance: the Problem of Sustained Attention*. London: Pergamon Press. 1971. Pp. 106. £2.95.

VIDERMAN, S. *La construction de l'espace analytique*. Paris: Editions Denoël. 1970. Pp. 348.

WELFORD, A. T. *Fundamentals of Skill*. London: Methuen. 1968. Pp. 426. £2.00.

WHITE, B. L. *Human Infants: Experience and Psychological Development*. Hemel Hempstead: Prentice-Hall. 1971. Pp. 160. £2.50.

YOUNG, J. Z. *An Introduction to the Study of Man*. Oxford: Clarendon Press. 1971. Pp. 719. £6.00.

Notes for Authors

The Journal publishes original papers on experimental work in all branches of human and animal psychology. Papers dealing with the mechanisms underlying behaviour are particularly welcome and theoretical papers on such topics are acceptable provided they relate to experimental findings. Short notes on apparatus likely to be of interest to experimental psychologists are also acceptable, provided that the design of the apparatus is an original contribution by the author.

Papers should be submitted in duplicate, one copy to be a *typewritten original,* preferably on A4 paper (210 × 297 mm), in double spacing with wide margins.

Title The title should be as concise as possible and should appear as a separate sheet together with the name(s) of the author(s) and the name of his institution. (The full postal address is not required on the title page.) A short running title of not more than 40 characters (including spaces) should be indicated if the full title is longer than this.

Abstract An abstract of 50–200 words should follow the title page as a separate sheet.

Headings and sub-headings for different sections of the paper should be clearly indicated. Authors should follow the conventions used in current issues of the Journal.

Acknowledgements should be as brief as possible and should appear as a separate sheet at the end of the text.

Footnotes should be avoided unless absolutely necessary. Essential footnotes should be indicated by superscript figures in the text and collected as a single page at the end of the manuscript.

References in the text should be indicated with the year of publication, in parentheses after the author's name, If more than one article by the same author in the same year is quoted, *a, b,* etc., should follow the year. When a paper has been published jointly by two authors, both names should be given in the text. For more than two authors, quote all names at the first mention and use the first author '*et al.*' for subsequent mentions. References to unpublished work should be avoided since they are of little use to readers. A list of the references quoted in the text should be given at the end of the paper in alphabetical order of author's names, commencing as a new sheet. Titles of Journals and books should be given *in full* thus:

WOODWORTH, R. S. (1901). On the voluntary control of the force of movement. *Psychological Review,* 8, 350–9.

GALTON, F. (1893). *Inquiries into Human Faculty and its Development.* London: Macmillan.

GARRETT, M. and FODOR, J. A. (1968). Psychological theories and linguistic construction. In DIXON, T. R. and HORTON, D. L. (Eds), *Verbal Behavior and General Behavior Theory.* Pp. 451–77. Englewood-Cliffs, N.J.: Prentice-Hall.

Tables should be kept to the minimum consistent with clarity of exposition. Each table should be typed double spaced as a separate sheet, giving the heading, e.g. TABLE II, in Roman numerals, followed by the legend, followed by the Table. Make sure that appropriate units are given. Instructions for placing the Table should be given in parentheses in the text, e.g. (TABLE II about here).

Figures are expensive to reproduce and should only be used when essential, The same data should not be presented both as a figure and in a table. Where possible, related diagrams should be grouped together to form a single figure. Figures should be drawn to professional standards in Indian Ink on Bristol Board of A4 size, and it is recommended that the linear dimensions of figures be approximately twice those intended for the final printed version. Lettering to appear in the figure should be done on tracing paper placed over the figure. Make sure that axes of graphs are properly labelled and that appropriate units are given.

The legend should be typed on a separate sheet of paper headed, e.g. FIGURE 2, in Arabic numerals. Instructions for placing the figure should be given in parentheses in the text, e.g. (FIGURE 2 about here).

Statistics Greek letters should be used to represent population parameters and Roman letters for sample values. Results of statistical tests should be given in the following form:

$$F = 25 \cdot 3, \, df = 1,9, \, P < 0 \cdot 001.$$

Similarly for t, χ^2 and other tests.

Units The Système Internationale (SI) will be used for all units. (If measures such as inches are used, then S.I. equivalents should be given.) If in doubt, refer to *Units, Symbols and Abbreviations*, published by the Royal Society of Medicine, 1 Wimpole Street, London, W.1.

The author will normally receive a decision on his paper within 3 months of receipt and if it is accepted it will normally be published 3 to 6 months later. The date of the receipt of the manuscript will be printed. Where minor revision of a paper is requested the original date of receipt will appear, provided that a satisfactory revision is received within 1 month of the request.

Contributors receive 25 copies of their paper free. Additional copies may be ordered on a form provided by the publishers at the time proofs are sent to authors.

Inhibition and Learning

edited by R. A. Boakes and M. S. Halliday

Laboratory of Experimental Psychology, University of Sussex, Sussex, England

April 1972
xvi + 568pp., £7·25

Inhibition and Learning arose from a conference held at the University of Sussex in Spring 1971 which brought together psychologists working in various traditions in an attempt to integrate recent theories and findings on the relationship between inhibition, learning and extinction. All the papers accepted for the book were thoroughly revised after the conference to take into account the points raised in discussion and the views put forward by other contributors. The result, a cohesive volume in which the approaches of the various traditions are compared and contrasted, will interest research workers in animal learning and physiological psychology, advanced undergraduate and graduate students of psychology.

Contents

Journal of Human Evolution

edited by A. B. Chiarelli

Institute of Anthropology Primatology Centre
University of Turin, Italy

This new journal will cover all aspects of human evolution including those concerned with non-human primates and will contain articles on the biochemical, skeletal and physiological development of Man. It attempts to integrate the many disciplines involved in the study of Man and his evolving behaviour and will reflect current interest in cytogenetics and heredity as well as paleontology and taxonomy. Problems of specificity will be treated alongside studies in biological variability, and anthropological aspects considered in social, cultural and physical contexts. It will appeal to libraries, museums and research institutes and meet the needs of students and scientists in many fields.

Contents of Volume 1 number 1

M. Masali: The postural relevance of Sergi's Temporal Tetrahedron in primates. A. Falek: Differential fertility and intelligence; resolution of two paradoxes and formation of a third. N. Petit-Marie: Evolution trends and comparative ontogenesis in primate cranium. M. G. Manfredi-Romanini: The nuclear DNA content and area of primate lymphocytes as a cytotaxonomical tool. M. L. Weiss: The frequency and maintenance of genetic variability in natural populations of *Macara fascicularis*. W. C. Osman Hill: Taxonomic status of the Macaques *Macaca mulatta* Zimm. and *Macaca Irus* Cuvier (= *M. fascicularis* Raffles). F. D. Burton and M. J. A. Bick: A drift in time can define a deme: the implications of tradition drift in primate societies for hominid evolution.

A. B. Chiarelli: New data on prosimiae chromosomes. A. A. Cigna: A consideration of modern dating methods for archaeological purposes. R. Fenart and R. Deblock: Ontogenic course of the teeth and the tooth-germs of anthropoids, studied in the vestibular axis. M. I. Satinoff: The origins and geographical spread of the thelassaemias and abnormal haemoglobins. L. E. M. de Boer: Marked chromosome associations in catarrhine monkeys, with a note on chromosome associations in other primate groups. D. Soudek: Chromosomal homozygous variants in Man. G. J. Armelagos, J. H. Mielke, K. H. Owen, D. P. Van Gerven, J. R. Dewey, and P. E. Mahler: Bone growth and development in prehistoric populations from Sudanese Nubia. Book reviews. Book news.

CONTENTS

OBITUARY NOTICE

J. ALFRED LEONARD, M.A. Ph.D.

Alfred Leonard came to Britain from Germany in 1938 and, like so many of the refugees of that time, he contributed in full to the life of his adopted country.

After the war, during which he was first interned and then served in the British Army, he took his degree in Psychology at Cambridge under Sir Frederic Bartlett and read for his Ph.D in the Nuffield Unit on Ageing. This period was the big formative influence of his academic life and his affection and respect for Sir Frederick never left him. Later, between 1954-6, he spent two years at Ohio State University working with Paul Fitts, from whom he learned a highly systematic approach to the organization of research and a determination to push it to the stage of practical implementation.

Alfred continued to work at the Medical Research Council's Applied Psychology Unit at Cambridge until 1965 and established an international reputation for his studies on skilled performance, choice reaction times and training. When he turned to the difficult problems of how the blind should be trained to move in an urban environment he was admirably equipped for the work which was to prove his most sustained contribution. It was a combination of his warm human sympathies and his determination never to lower his high scientific standards. In his hands the most hard-headed and quantitative experiments were at the same time imaginative and sensitive to the needs of blind people. He joined the Department of Psychology at Nottingham in 1965 where, with support primarily from the Medical Research Council, he was able to expand the range of his work. He contributed notably to developing new methods of assessing the needs of blind people, to the relative effectiveness of different mobility aids and to training methods, in particular those associated with the use of the long cane.

Alfred Leonard's untimely death prevented him from completing the work he had planned but we all recognize the amount he achieved in such a short time. It is the monument of a man who was always a crusader and who demonstrated in a very direct way how science can be made to serve human purposes and how a humane concern can improve the quality of science.

C.I.H.
H.K.

ENCODING AND MEMORY FOR VERBAL AND PICTORIAL STIMULI

J. ELISABETH WELLS

University of Canterbury, New Zealand†

Memory for visually presented verbal and pictorial material was compared using stimuli chosen to minimize non-essential differences between the two types of material. Experiment I required retention of a short list; verbal and pictorial stimuli were remembered equally well. Experiment II required recall of single items after 30 s of backwards counting; recall was much superior for pictorial stimuli. The type of task appeared to affect encoding, with verbal encoding reported to be predominant in Experiment I and visual encoding, or imagery, common in Experiment II.

Introduction

During the last decade there have been a number of experiments comparing memory for verbal and pictorial material (see Paivio, 1969). Such studies attempt to present the same information in two ways, verbally and pictorially, to see if the mode of presentation has any effect on memory. In other words, are nameable pictorial stimuli remembered better or worse than their verbal equivalents? Characteristically, objects, photographs, or drawings have been used as pictorial stimuli, with single noun labels as the verbal equivalents of the pictorial stimuli. The purpose of this paper is, first of all, to show that conclusions drawn from experiments using such stimuli may be misleading, and, second, to report two experiments using stimuli chosen to remove the confounding present in earlier experiments.

Experimenters (e.g. Paivio and Yarmey, 1966) have been careful to ensure that subjects would label the pictorial stimuli with the words used as verbal stimuli, but little attention has been paid to other aspects of verbal/pictorial equivalence. In particular, the exemplar/class relationship of pictures and nouns has been virtually ignored, although it is obvious that, for instance, a chair or a drawing of a chair is a specific example of the class of objects called "chairs". One consequence of this exemplar/class relationship is that, because exemplars always possess characteristics additional to those defining the class, the number of cues available in a pictorial stimulus is always greater than the number necessarily implied by its noun label. However, we cannot conclude that, because more cues are available in pictures than in nouns, subjects will always encode more cues from pictures than from nouns. For example, a subject may encode only a subset of the cues available in a picture or object. Alternatively, when presented with a noun, he is free to imagine a referent with many supplementary characteristics. To summarize: experiments which compare memory for nouns and pictures (or nouns and objects) confound verbal/pictorial differences with class/exemplar differences. The effect of this confounding

---

† Now at the University of Toronto.

on the number of cues encoded from each type of stimulus cannot be precisely determined.

What kind of verbal and pictorial stimuli minimize class/exemplar differences and provide greater experimental control over encoding of cues? Dallett and Wilcox (1968) removed much of the class/exemplar distinction by providing paragraph descriptions of pictures, but the complexity of their pictures and descriptions resulted in even less control over the encoding of cues than in most other studies. Presentation of simple stimuli varying only on a few attributes seems to provide the most effective means of controlling encoding. Suppose that the pictorial stimuli vary only in size, colour, and pattern, and that the verbal descriptions corresponding to these stimuli consist of a size word, a colour word, and a pattern word, e.g. "large red dotted". Subjects are required to remember the size, colour, and pattern conveyed by each stimulus, whether verbal or pictorial. Presumably they will encode only size, colour and pattern, and perhaps some of the characteristics intrinsically linked with the type of stimulus, such as the type case of the words. With this sort of simple stimulus material, non-essential differences between verbal and pictorial stimuli appear to be minimized.

The two experiments reported here used these simple kinds of stimuli. The first experiment employed a recognition task to compare memory for short lists of verbal and pictorial stimuli, and the second experiment utilized a Peterson and Peterson (1959) type of task. One unavoidable difference between geometric stimuli and the words describing them is that, while size, colour and pattern can be presented on one compact blob, the words describing such a blob must be separate and sequential. Is this difference important? To obtain a partial answer to this question both experiments investigated the effects of spatial distribution on verbal stimuli and on pictorial stimuli. Obviously a verbal stimulus can never be as "compact" as a pictorial stimulus, but relatively compact and relatively distributed forms of each type of stimulus can be devised.

Some experimenters (e.g. Jenkins, Neale and Deno, 1967) have inferred that, if pictures are remembered better than words this will be because the pictures tend to be stored in both visual and verbal codes whereas words tend to be stored only in verbal codes. Corroborating evidence has been provided by Paivio and Csapo (1969) who found that inferences from identification, naming, and image arousal latencies agreed with inferences from memory data. In the following two experiments, encoding was investigated by means of a questionnaire. This procedure does not imply a belief that subjects would know how they encoded stimuli—it was merely an attempt to see if their responses would relate meaningfully to the memory data.

Experiment I
Method
Stimulus materials

As illustrated in Fig. 1, the stimuli were of four kinds: verbal compact, verbal distributed, pictorial compact, and pictorial distributed. Each set of stimulus material varied on three attributes (size, colour and pattern) with three levels per attribute (large, medium, small; red, blue, green; dotted, striped, checked). Compact pictorial stimuli presented all attributes on one blob whereas distributed pictorial stimuli had three blobs, the left one conveying size

(a) Verbal compact (b) Verbal distributed

(c) Pictorial compact (d) Pictorial distributed
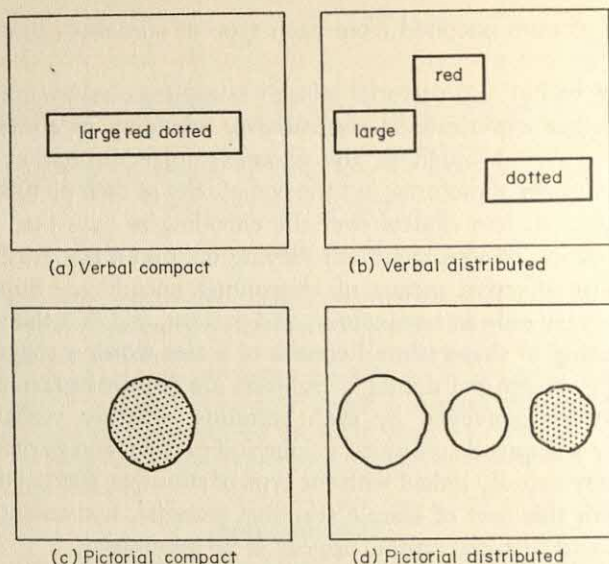
FIGURE 1. The four kinds of stimuli.

information, the middle one colour, and the right one pattern. For verbal stimuli the compact form presented three words in one central box and the distributed form showed three boxes with one word in each. The colour of the pictorial stimuli is, of course, not shown in Fig. 1.

Recognition templates

Each template contained a complete set of 27 stimuli. Because 27 distributed stimuli could not be fitted on standard size 34 × 21 cm sheets without overcrowding, only compact stimuli were used. This resulted in two kinds of recognition templates, verbal compact, and pictorial compact.

Subjects

The subjects were undergraduate psychology students completing course requirements at the University of Canterbury.

Design

The four between-subject factors in the 2 × 2 × 2 × 2 design were stimulus type (verbal/pictorial), stimulus distribution (compact/distributed), recognition template type (verbal compact/pictorial compact), and sex (male/female subjects). With 6 subjects in each of the 16 cells, there were a total of 96 subjects.

Procedure

The 35 mm stimulus slides were projected on to a 1·2 × 1·8 m screen by a Leitz 500 W projector with a mechanical shutter, and each slide was shown for 0·7 s with an inter-slide interval of 0·5 s. On each trial three stimulus slides were presented sequentially and then subjects had to mark on their recognition templates the three stimuli they had just seen. Response order was free but subjects were required to mark "1" for the first slide they had seen, "2" for the second, and "3" for the third. They were given two practice trials with stimuli of the same type as in the experimental trials (but with different colours and patterns) and 54 experimental trials.

On each trial, no two temporally adjacent stimuli shared a common attribute level. For example, the first and third stimuli might both be red but the first and second or the second and third would never both be red. Apart from this restriction, the three stimuli for each trial were chosen randomly.

Questionnaire

After the experimental trials, all subjects filled out the same questionnaire. Only the questions on encoding are of interest here.

(a) Did you, after seeing a stimulus (i.e. a slide) on the screen, think of it in words e.g. "large red dotted"?

(b) If you did think of stimuli in words at all, did you "hear" these words in your mind?

(c) Did you "see" the stimuli in your mind, either as words or as pictures?

Subjects replied to each question by circling "Always", "Often", "Sometimes", or "Never".

Control group

A control group of 24 subjects was run to check that all types of stimuli were clearly visible. There were three males and three females in each of the four stimulus conditions. After viewing the practice stimuli, subjects were shown each of the 27 stimuli twice, in random order, making a total of 54 trials. On each trial a stimulus slide was presented for 0·7 s and subjects wrote down the size, colour, and pattern. The overall error rate was very low (0·86%) and results for all groups were similar.

Results

Recognition task

Responses could be scored in several ways but, as these measures were highly correlated, the main analyses were carried out only on the number of attributes correct when serial position was reported correctly. In the absence of any interaction between compact/distributedness and verbal/pictorial stimulus type, compact/distributedness was treated as the same factor for verbal stimuli as it was for pictorial stimuli.

The mean per cent correct recognition for each condition is shown in Table I.

TABLE I

Experiment I. Mean per cent correct recognition

		Verbal stimuli		Pictorial stimuli	
		Compact	Distributed	Compact	Distributed
Verbal	Males	84·1	74·0	73·9	69·5
Recognition	Females	77·5	75·1	86·1	77·3
Pictorial	Males	82·2	75·2	83·8	78·3
Recognition	Females	85·9	81·1	83·5	77·1

A four-way ANOVA showed compact/distributedness to produce the only significant main effect, $F = 8·30$, $df = 1, 80$, $P < 0·01$, compact stimuli being recognized better than distributed stimuli. The verbal/pictorial stimulus factor did not even approach significance, $F = 0·10$, $df = 1, 80$, $P > 0·25$. There was a just significant three-way interaction, $F = 4·59$, $df = 1, 80$, $P < 0·05$, which suggested, if anything, that males found mixed modes (e.g. verbal stimuli but pictorial recognition templates) more difficult than single modes whereas the reverse was true for females. The significant main effect and interaction accounted for little of the variance—about 10%.

The attribute scores used in the above ANOVA were obtained by summing over trials, serial position within a trial, and size, colour, and pattern. Following the example of Murdock and Ogilvie (1968), these attribute scores were broken down into size, colour, and pattern scores and another ANOVA was carried out with the type of attribute as a fifth factor. The main effect of attribute type was significant beyond the o·oo1 level, $F = 19·5$, $df = 2$, 60, size being remembered better than colour, which was remembered much better than pattern. The attribute scores were also separately broken down into scores for each serial position within a trial. An ANOVA with position as the fifth factor showed an extremely large main effect, $F = 69·3$, $df = 2$, 160, $P < 0·001$, with the first and third stimuli being remembered much better than the middle one.

Questionnaire

The responses of Verbal Stimulus subjects were compared with those of Pictorial Stimulus subjects for each of the three questions on encoding. The results were all non-significant—$\chi^2 = 0·8$, 0·6, and 5·3 with $df = 2, 2,$ and 3, for questions (a), (b), and (c) respectively. Similar comparisons between the two levels of each of the other experimental factors were likewise all non-significant.

For each question on encoding, one-way ANOVAs were used to compare the recognition scores of subjects who responded "Always" with the scores of subjects who responded "Sometimes", or "Often", or "Never". As Table II shows, the more "auditory" encoding reported [question (b)], the better the recognition scores, and a similar but not quite significant trend occurred with verbal encoding [question (a)]. Responses to question (c), however, showed no relation at all between the amount of reported visual imagery and recognition scores.

TABLE II

Experiment I. Mean per cent correct recognition at each level of encoding

Question	Always	Level of encoding Sometimes	Often	Never	F	df
(a)	81·6	78·2	73·7	†	2·89	2,93
(b)	84·0	76·8	74·9	73·6	5·35‡	3,92
(c)	80·8	80·6	76·7	80·7	1·07	3,92

† No observations.
‡ $P < 0·01$.

Experiment II

Contrary to most previous experiments, Experiment I showed that verbal and pictorial stimuli were remembered equally well and the questionnaire data suggested no differences in the encoding of the two types of stimuli. Only verbal encoding [questions (a) and (b)] related to better recognition; visual encoding, or imagery as it might be called, appeared to deserve Neisser's (1967, p. 157) appellation of a "cognitive luxury".

A second experiment was carried out to see if the results of Experiment I were

specific to the task used. Would pictorial and verbal stimuli always be remembered equally well when non-essential differences between the two types of stimuli had been minimized? This seemed unlikely in view of a number of studies showing different results for different tasks. For example, Paivio and Csapo (1969) found pictures to be remembered better than words in free recall, but not in memory span, with the results in serial recall dependant on the rate of presentation. Their hypothesis that imagery was important only in non-sequential tasks, suggested that imagery might be more important in memory for single items instead of the series of three items used in Experiment I. The Peterson and Peterson (1959) technique of item presentation, interpolated activity, and recall, seemed particularly suitable. Recall was used instead of recognition in Experiment II, because it was suspected that the compact/distributed effect in Experiment I was the result of presenting only compact stimuli on the recognition templates. Thus subjects had to convert remembered distributed items to compact form before responding.

Because the amount of reported "auditory" encoding in Experiment I related to better recognition, an attempt was made in Experiment II to manipulate such encoding experimentally by introducing a factor of silence or voicing at presentation.

Method

Stimulus materials

These were the stimuli in Experiment I.

Subjects

The subjects were undergraduate psychology students completing course requirements at the University of Canterbury.

Design

This was a $2 \times 2 \times 2 \times 2 \times 2$ factorial design with 6 subjects per cell. The five between-subjects factors were stimulus type (verbal/pictorial), stimulus distribution (compact/distributed), vocalization (silence/voicing), sex, and retention interval (0 or 30 s). Actually each subject occupied a unique cell in the total design because there were two additional between-subjects factors—three presentation orders of stimuli and two sets of numbers for backwards counting. Presentation order and number set were not included as factors in the analyses of variance.

Procedure

All subjects were run individually. A subject was first given two 30 s periods of practice at counting backwards by threes to the 1 s beat of a metronome. Then he viewed three practice stimuli for 5 s each, to familiarize him with the type of stimuli he would see. A fourth practice stimulus slide was shown for 0·8 s, the time of presentation in the experimental trials. Then a chart showing the sizes, colours, and patterns in the stimuli was studied for 1–2 min. There were nine experimental trials and on each one a stimulus to be remembered was shown for 0·8 s, followed by a number slide which was shown for 0·7 s after a 1·0 s inter-slide interval. The subject counted backwards for 30 s and then had 20 s for written recall (in words). Auditory signals were used to indicate recall and the onset of the next trial.

The procedure was identical for the immediate recall groups except that the order of backwards counting and recall was reversed. Recall began 1 s after the stimulus had been presented, then 20 s later the number slide was presented, and subject counted backwards for 30 s.

Questionnaire

After the experimental trials all subjects, except those in the immediate recall groups, filled out the same questionnaire. The questions on encoding were:

When you *saw* the slide you had to remember

(a) did you think of it in words, e.g. "large red dotted"?
(b) if you did think of a slide in words at all, did you "hear" these words in your mind?

When trying *to remember* the slide you had seen

(c) did you remember it in words?
(d) did you seem to "hear" what it was?
(e) did you see an image of words?
(f) did you see an image of a picture?

Subjects replied to each question by circling "Always", "Often", "Sometimes", or "Never".

Results

Recall task

All groups were scored for attribute errors, e.g. if, for any item, only size was recalled incorrectly, then that item would be scored as 1 wrong and 2 right (colour and pattern).

As in Experiment I, the immediate recall (0 s retention) groups showed perceptual errors to be rare—the error rate was 1·35%—and χ^2 tests on the number of errors per subject revealed no differences between groups. Results from the 0 and the 30 s retention groups could not be analysed together as the range of errors was so very different.

Results from the ANOVA for the 30 s groups were in sharp contrast to those from Experiment I; pictorial stimuli were recalled much better than words, $F = 29·2$, $df = 1, 80$, $P < 0·001$, accounting for 20% of the variance, compact/distributedness and voicing had no effect, $F = 3·10$ and $0·99$, $df = 1, 80$, respectively, and females did slightly better than males, $F = 5·16$, $df = 1, 80$, $P < 0·05$. Mean error scores for each condition are shown in Table III.

TABLE III

Experiment II. Mean number of errors in the 30 s retention groups

		Verbal stimuli		Pictorial stimuli	
		Compact	Distributed	Compact	Distributed
Silence	Males	6·83	8·67	4·50	4·00
	Females	4·50	6·00	3·17	3·50
Voicing	Males	7·83	10·17	3·00	4·67
	Females	6·83	7·00	2·50	4·33

Another ANOVA was carried out with the attribute errors partitioned into size, colour, and pattern, as a fifth factor; pattern was recalled better than size or colour, $F = 3·83$, $df = 2, 160$, $P < 0·05$, but this order of difficulty was confounded with order of report. Most subjects reported size, colour, then pattern but some reported colour before size. Attribute errors were also separately broken down into errors per trial and the five-way ANOVA showed trials to have a significant effect,

$F = 2 \cdot 99$, $df = 8, 640$, $P < 0 \cdot 01$. The curve was bow shaped, beginning at $85 \cdot 4\%$ of attributes correct, falling to $71 \cdot 5\%$ by the fourth trial, and rising again to $81 \cdot 6\%$ correct on the ninth trial.

Questionnaire

Of the experimental variables of verbal/pictorial stimulus type, compact/distributedness, voicing, and sex, only stimulus type affected answers to any of the questions on encoding. The results are shown in Table IV. Verbal encoding, as

TABLE IV

Experiment II. Encoding of verbal and pictorial stimuli

Question	Relation	x^2	df
(a)	V = P	6·2	3
(b)	V = P	2·4	3
(c)	V = P	1·4	3
(d)	V = P	1·8	2
(e)	V > P	11·6†	2
(f)	V < P	27·5‡	3

Note—V indicates the amount of encoding reported by subjects who saw verbal stimuli; P indicates the amount of encoding reported by subjects who saw pictorial stimuli.
† $P < 0 \cdot 01$.
‡ $P < 0 \cdot 001$.

indicated by responses to question (a), (b), (c) and (d), occurred an equal amount for both verbal and pictorial stimuli. However, pictorial stimuli were much more likely than verbal stimuli to be recalled as visual images of pictures [question (f)]. Nonetheless, 20 out of the 48 subjects who saw words reported at least sometimes experiencing images of pictures. Not surprisingly, images of words during recall were more common for verbal stimuli than for pictorial stimuli [question (e)].

Of the six questions on encoding, only responses to (d) and (f) related to recall scores. The more that recall in "heard" words was reported (d), the less errors were made ($P < 0 \cdot 001$, using Wilkinson's (1951) tables for evaluating a series of t-tests). Likewise, the more that subjects reported pictorial images at recall, the better their recall was ($P < 0 \cdot 005$, using Wilkinson's tables). The mean number of errors at each level of encoding is given in Table V.

TABLE V

Experiment II. Mean number of errors at each level of encoding

Question	Always	Level of encoding Often	Sometimes	Never
(a)	2·8	3·8	6·2	6·3
(f)	3·8	4·6	4·7	7·4

Assuming that responses to questions (d) and (f) reflect the codes in which stimuli were stored, it can be seen that while verbal and pictorial stimuli were encoded verbally to the same extent, pictorial stimuli were more often *also* encoded in some visual, pictorial form. Thus, when pictorial stimuli were remembered better than verbal stimuli, the "summative encoding" (Paivio and Csapo, 1969) of the pictorial stimuli was greater than that of the verbal stimuli.

Discussion

A comparison of Experiments I and II shows that the type of task affects memory for verbal and pictorial material. It seems unlikely, though, that this is because the first experiment involved recognition, and the second, recall. In both experiments subjects were aware of the complete set of combinations of size, colour, and pattern which could occur, and Dale (1967) has shown that when subjects know all the response alternatives there is no difference between recognition and recall.

Two factors may be postulated to account for the differences in the results from the two experiments. The first factor is the serial nature of the task. Experiment I involved memory for a series of three triple-attribute items; response order was free but subjects had to indicate the presentation order of the three items. In Experiment II, however, there was only one triple-attribute item to be remembered on each trial. The second factor is the type of interference in each task. In Experiment I each item was preceded or followed by items of the same type, whereas in Experiment II each item was followed by backwards counting, the interpolated activity.

The effect of the serial factor can be predicted from the theory, expounded most clearly by Paivio and Csapo (1969), which assumes that only verbal encoding is important for sequential memory whereas both verbal encoding and pictorial imagery are important for non-sequential memory. These assumptions were supported by the finding that in Experiment I, only verbal encoding related to better performance, whereas in Experiment II both verbal encoding and concrete visual imagery related to superior recall. Second, given that the pictorial stimuli can readily be named by subjects, the theory predicts that (a) in a serial task, both verbal and pictorial stimuli will be encoded verbally and so both types of stimuli will be remembered equally well; and (b) in a non-sequential task, both types of stimuli will be encoded verbally but pictorial stimuli will be more likely than verbal stimuli to be also encoded as concrete visual images—and as performance in such a task depends on both verbal and imagery codes, pictorial stimuli will be remembered better than verbal stimuli. As predicted, verbal and pictorial stimuli were remembered equally well in Experiment I whereas, in Experiment II, pictorial stimuli were remembered much better than verbal stimuli. The questionnaire data showed that, also as predicted, verbal and pictorial stimuli were encoded the same way in Experiment I, but in Experiment II imagery was more common with pictorial stimuli, although both types of stimuli were verbally encoded to the same extent.

The results obtained from Experiments I and II are also what would be expected from consideration of the type of interference in each task. In Experiment I

interference came from other items of the same type. As each type of stimuli varied only in the information conveyed about size, colour and pattern, the intra-list similarity of three pictorial stimuli should be the same as that of three verbal stimuli. Therefore the pictorial stimuli could be expected to have interfered with other pictorial stimuli to the same extent as verbal stimuli interfered with other verbal stimuli. Thus words and pictures would be remembered equally well, as indeed they were. This result should obtain, regardless of how the stimuli are encoded, as long as stimuli of the same type are encoded in the same way. However, if memory for three triple-attribute stimuli involved a rehearsal buffer, which is usually assumed to be verbal, then it is reasonable to expect verbal encoding for both types of stimuli.

In Experiment II the main source of interference was the verbal interpolated activity of backwards counting. If interference is specific to the type of encoding used (Cohen and Granström, 1970), this interpolated activity should have interfered with verbally encoded information more than with visually encoded information. As pictorial stimuli are more likely to be stored visually than are verbal stimuli, even allowing for imagery, the pictorial stimuli should have been less affected by the backwards counting and therefore should have been recalled better than the verbal stimuli. The data confirmed this prediction.

These two factors, the serial nature of the task and the type of interference, presumably both acted to maximize the differences in results between the experiments. Consideration of each factor leads to the prediction that subjects will remember verbal and pictorial stimuli equally well in Experiment I, whereas in Experiment II pictorial stimuli will be remembered better. These predictions were borne out by the memory data. Inferences about encoding from the memory data and evidence from that rather weak source, the questionnaire data, confirmed the predictions that stimuli in the first experiment would be verbally encoded, although in the second experiment visual encoding or imagery, would be important.

Minimizing non-essential differences between pictorial and verbal stimuli does not necessarily result in both types of stimuli being remembered equally well. This suggests that previous results showing superior memory for pictures were probably not an artifact, caused by the greater number of cues conveyed by the pictures than by the words. Whether or not nameable pictorial stimuli will be remembered better than their verbal equivalents depends on the task. The type of task appears to affect the encoding used, and it seems that in those cases where visual encoding is common, pictorial stimuli are remembered better. The results of the two experiments reported in this paper are in agreement with Tversky's (1969) finding that not only can pictorial stimuli be encoded verbally and vice versa, but that the type of encoding used depends on the task.

References

COHEN, R. L. and GRANSTRÖM, K. (1970). Reproduction and recognition in short-term visual memory. *Quarterly Journal of Experimental Psychology*, 22, 450–7.

DALE, H. C. A. (1967). Familiarity and free recall. *Quarterly Journal of Experimental Psychology*, 19, 103–8.

DALLETT, K. and WILCOX, S. G. (1968). Remembering pictures versus remembering descriptions. *Psychonomic Science*, 11, 139–40.

JENKINS, J. R., NEALE, D. C. and DENO, S. L. (1967).   Differential memory for picture and word stimuli. *Journal of Educational Psychology*, **58**, 303–7.

MURDOCK, B. B. JR. and OGILVIE, J. C. (1968).   Binomial variability in short-term memory. *Psychological Bulletin*, **70**, 256–60.

NEISSER, U. (1967).   *Cognitive Psychology*.   New York: Appleton-Century-Crofts.

PAIVIO, A. (1969).   Mental imagery in associative learning and memory. *Psychological Review*, **76**, 241–63.

PAIVIO, A. and CSAPO, K. (1969).   Concrete image and verbal memory codes. *Journal of Experimental Psychology*, **80**, 279–85.

PAIVIO, A. and YARMEY, A. D. (1966).   Pictures versus words as stimuli and responses in paired-associate learning. *Psychonomic Science*, **5**, 235–6.

PETERSON, L. R. and PETERSON, M. J. (1959).   Short-term retention of individual verbal items. *Journal of Experimental Psychology*, **58**, 193–8.

SHEPARD, R. N., HOVLAND, C. I. and JENKINS, H. M. (1961).   Learning and memorization of classification. *Psychological Monographs*, **75**, (13, Whole No. 517).

TVERSKY, B. (1969).   Pictorial and verbal encoding in a short term memory task. *Perception and Psychophysics*, **6**, 225–33.

WILKINSON, B. (1951).   A statistical consideration in psychological research. *Psychological Bulletin*, **48**, 156–8.

SHORT-TERM RETENTION OF TACTILE STIMULATION

EDITH V. SULLIVAN AND M. T. TURVEY

Department of Psychology, University of Connecticut

In three experiments subjects were required to reproduce after varying delays the locus of a tactile stimulation delivered to the upper-side of the arm. During the retention periods subjects either performed a subsidiary, arithmetic task or rested. Recall, as measured by accuracy in reproducing the locus of stimulation, decreased as a function of retention interval, asymptoting after approximately 5 s. Performance was poorer in the subsidiary task condition than in the rest condition; however, the effect of the subsidiary task appeared to be more on subject recall strategies than on rehearsal capacity. No evidence of proactive interference effects was found, and a decay interpretation of forgetting of discrete tactile stimuli in the short-term memory distractor paradigm was favoured.

Introduction

Recently a tactile analogue to verbal short-term memory (STM) was investigated by Gilson and Baddeley (1969). These investigators employed the Brown-Peterson distractor paradigm (Brown, 1958; Peterson and Peterson, 1959) with the locus of a stimulation to the underside of the arm as the to-be-remembered material. The ability of subjects to reproduce the locus of the tactile stimulation was investigated as a function of length of retention interval and retention interval activity in a between-subject design. During the retention interval the subject was either permitted to rest (R) or required to perform a subsidiary task (ST) of counting backwards by threes from a three-digit number. The Gilson and Baddeley experiment revealed no difference between the R and ST conditions across the first 10 s; beyond 10 s, however, there was a very large decrement in the performance of the ST group compared to the performance of the R group. Gilson and Baddeley interpreted their results as indicating a dual memory process for tactile stimuli, a rapidly decaying trace independent of rehearsal capacity and a subsequent, more persistent trace, dependent on rehearsal capacity.

These results are not in accord with STM data for verbal material obtained in equivalent experimental situations. Dillon and Reid (1969) have demonstrated that the effect of a rehearsal-preventing ST on verbal STM is more pronounced in the early part of the retention interval than in the later part. Furthermore, the results of Gilson and Baddeley are at odds with reports of the effects of ST on STM for motor acts. Whereas increasing the time spent on ST, up to a point, increases the forgetting in verbal STM, the effect of ST on motor STM tends to produce a reduction in recall accuracy which is relatively independent of the time spent on ST (Pepper and Herman, 1969).

In part the present series of experiments is a replication of Gilson's and Baddeley's experiment. Tactile STM was examined in the distractor STM

paradigm as a function of R, ST and retention interval. Also examined was the question of whether tactile forgetting in the distractor STM situation is due to proactive interference (PI).

General Method and Apparatus

The apparatus for delivering the tactile stimuli was constructed with a smooth board, 12 × 6 × 1 in (1 in = 25·4 mm), as the base. Attached to one side of the board were two L-shaped metal brackets extending upwards. The brackets supported a slide rule with the centre section removed. Through the centre line of the slider, a hole of about 1/32 in was made through which a pointer providing the stimulation was inserted. The calibrations on the slide rule provided experimental measures accurate to 1/50 in.

Each subject positioned her left arm on the board under the slide rule so that the top part of the forearm was directed upwards. Gilson and Baddeley used the underside of the arm. Right-handed females served as subjects for all three experiments to minimize possible confounding resulting from handedness, laterality or sex factors (Weinstein and Sersen, 1961). The subjects were run individually by two female experimenters and received all experimental conditions in one session. The subject's arm was secured to one position for the duration of the experiment by two canvas straps, one at the wrist bone and the other near the elbow. For all experiments, stimulus points were centrally located on the subject's arm. During no part of the actual experiments were subjects permitted to view the apparatus. The experimenter delivered the stimuli via the sliding pointer, and subjects recalled the locations of stimulation via the same procedure. Between stimulation and reports, the pointer was always returned to the wrist-end of the apparatus. In the Gilson and Baddeley experiment, the experimenter used the tip of a pen as the stimulator, and the subject reported the locus of stimulation by pointing with the contralateral hand.

For each experiment of the present study, subjects received some degree of training on the tactile STM task. Recall scores were absolute deviations in inches of the recall points from the stimulus points.

Experiment I

Subjects

Sixteen right-handed female students from introductory psychology courses at the University of Connecticut served as subjects. Participation as subjects was part of their required course work.

Apparatus

The apparatus described above was employed. The subjects sat at a 45° angle to a table on which the apparatus was situated.

Procedure

A within-subject design was used. Gilson and Baddeley had used a between-subject design. Pilot work, however, had pointed to considerable variation in accuracy on the tactile STM task, favouring, therefore, the usage of a within-subject design. A Brown-Peterson STM procedure served as the experimental paradigm.

At the beginning of the experiment, subjects received either R or ST instructions. For R, subjects were told to rehearse the point of stimulation "in their minds" during the retention interval. For ST, subjects were instructed to write down with their right hands the sums of 5 pairs of digits and the classifications of those sums as odd or even. Digit pairs were printed in columns on a sheet of paper which was on the arm rest to the right of the subject. The subjects were told that immediately after feeling the stimulus they should begin to add and classify the digit pairs, completing as many as possible during the retention period.

There were four retention periods, 0, 2, 10 and 30 s. Each retention period was examined under R and ST instructions. Five trials were given for each retention interval–retention

interval activity combination. A different locus of stimulation was given on each trial. In all, 20 loci within a 3·80 in range were used with no locus being repeated within the four blocks of five trials which comprised a subject's testing under one of the instruction conditions. A quasi-randomization of stimuli was used with the stipulation that adjacent loci would not be allowed within any given time condition. Thus, likely proximity and confusion effects were minimized. The subject was examined first under one instruction set, then under the other. Upon completing a block of five trials at a particular interval, the subject was informed that there would be a change in the interval. After completing the 4 blocks of five trials under, say, the R condition, the subject was then given the other set of instructions. At this juncture the subject's arm was removed from the apparatus to prevent fatigue. The order in which the retention periods were tested within an instruction condition and the order in which the instructions were given were counterbalanced across subjects.

On a trial the subject was cued to recall by the experimenter's saying "recall" after the predetermined retention interval. Several seconds elapsed between the recall cue and the subject's attempt to reproduce the locus of stimulation. This was the time required by the subject to move her right hand to the pointer of the apparatus. The o s retention interval, therefore, represents at best the shortest delay at which retention could be tested.

Results and Discussion

Mean deviation in inches was calculated for each block of five trials for each subject. The results are represented in Figure 1. A Treatment by Subjects
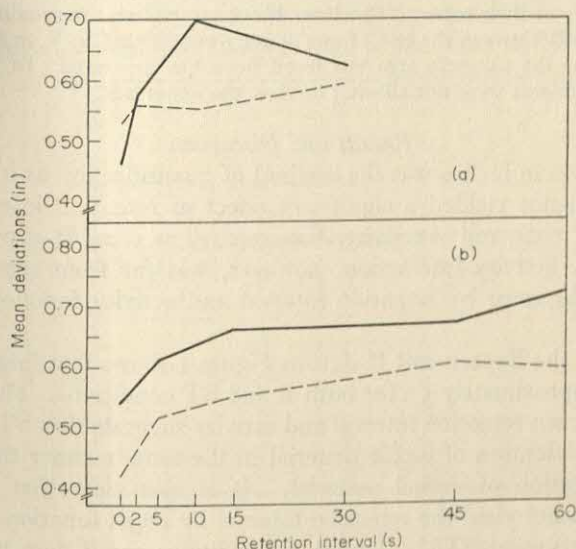
FIGURE 1. Short-term retention as a function of retention interval with type of subsidiary task as the curve parameter observed in Experiment 1 (a) and in Experiment II (b). ——ST; - - - - R.

analysis of variance revealed that neither the main effect of retention interval, $F = 2·47$, $df = 3, 45$, $0·05 < P < 0·10$, nor the main effect of retention interval activity, $F < 1$, was significant. The interaction between retention interval and activity also failed to achieve significance, $F = 1·87$, $df = 3, 45$, $P > 0·05$.

Experiment II

Despite the absence of significant main effects, the data of Experiment I clearly suggested caution in accepting the null hypothesis. This caution was further

supported by the following considerations: First, in the Gilson and Baddeley experiment the effects of retention interval and activity became more evident at longer retention intervals somewhat in excess of those used in Experiment I. Second, in retrospect, the procedure used in Experiment I was found wanting on two counts: the amount of time allotted to recall should have been controlled across subjects, and also it was clearly necessary to emphasize to subjects the need to search for the precise location of the tactile stimulus if they were slightly uncertain about the accuracy of their initial recalls. Experiment II was designed with those considerations in mind.

Procedure

A within-subject design was used with 25 subjects drawn from the same population from which the subjects of Experiment I had been drawn. The general procedure and instructions of Experiment I were employed with certain modifications. The retention intervals in Experiment II were 0, 5, 15, 45 and 60 s. For each interval-activity (R or ST) combination, the subject received five trials. The locus of the tactile stimulus varied on each trial of a block. The total of 25 loci were used for the five blocks of trials comprising an activity condition. The loci were within a 4·8 in range, and the average distance between loci was 0·20 in. The subjects were required to recall within 15 s after the recall cue. The subjects were instructed to move the pointer within this 15 s period until they felt confident about the accuracy of their report. The inter-block interval was approximately 1 min, and about 5 min elapsed between the end of one activity condition (i.e. R or ST) and the next during which time the subject's arm was freed from the apparatus. In the course of the experiment the subjects were not allowed to view the apparatus.

Results and Discussion

Mean deviation in inches was the method of quantification as in Experiment I. Analysis of variance yielded a significant effect of retention interval, $F = 3·40$, $df = 4, 96, P < 0·05$, and of activity, $F = 5·40, df = 1, 24, P < 0·05$. The retention interval by activity interaction, however, was far from significant, $F < 1$, $df = 4, 96$. The error by retention interval and activity functions are given in Figure 1.

Inspection of the Experiment II data in Figure 1 shows that forgetting achieved asymptote at approximately 5 s for both R and ST conditions. The absence of an interaction between retention interval and activity suggests that ST does not affect the short-term retention of tactile material in the same manner that it affects the short-term retention of verbal material. It is also clear that opportunity to "rehearse" does not yield the retention interval by recall function in tactile STM that it yields in verbal STM. Under R conditions recall does not decline as a function of retention interval in verbal STM situations. Experiment II, however, showed that given the opportunity to rehearse, i.e. unoccupied retention periods, forgetting still occurred for tactile material.

The difference between the R and ST conditions at 0 s is somewhat puzzling. When a subject performed the 0-s block of trials under ST instructions she did not of course perform the ST. Therefore, the 0-s data points should have been identical for R and ST. The subjects, however, not only had a cognitive set to perform ST in the ST 0-s condition, but they indeed attempted to pick up the pencil available for the performing of ST. Since subjects were not informed as to how long the retention interval for a block of five trials was to be, this set to perform

ST accompanied by an attempt to perform ST, might have affected level of retention, at least on the first trial of a o-s data block. The o-s points were recalculated. This time, the first trial was not included in the calculations. Results showed that instead of narrowing the gap in question, the difference tended to widen. The ST o-s data point calculated from all five trials was 0·54; calculated from the last four trials it was 0·55. The R o-s data point based on all five trials was 0·42; calculated from the last four trials it was 0·37.

If ST were acting as a true source of rehearsal prevention, then increasing the time spent on ST should, presumably, have increased the degree of separation between R and ST functions. This, however, did not occur. Therefore, an explanation of the R-ST difference was sought elsewhere.

The R data of Experiment II were partitioned into the data for those subjects who performed R before ST and the data of those subjects who performed R subsequent to ST. A similar partitioning was conducted on the ST data. The four resulting functions are shown in Figure 2. An analysis of variance was conducted separately on R vs. ST as first-performed conditions, and R vs. ST as second-performed conditions. The time by activity interactions were not significant, either for ST-R first or for ST-R second conditions alone; F was less than one in both cases. There was an overall sequence by activity interaction ($F = 4·47, df = 1, 44, P < 0·05$). Separate analyses showed that in the case of ST-R first, no real
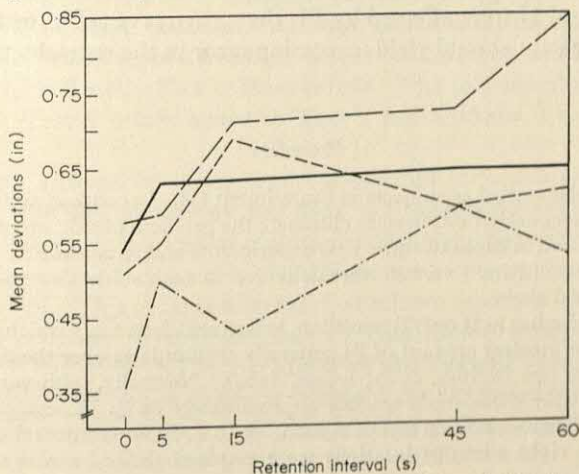
FIGURE 2. Short-term retention as a function of retention interval with the order of subsidiary task as the curve parameter observed in Experiment II. ——— ST first; - - - - R first; — — — ST second; ·—·—·— R second.

differences were found; in fact, only the time condition resulted in an F greater than one. When ST or R were given second, both time ($F = 3·03, df = 4, 88, P < 0·025$) and activity ($F = 5·48, df = 1, 22, P < 0·025$) were significant, thus suggesting that the weight of the overall difference could be laid on the second part of each experimental condition.

The R-ST difference represented in Figure 1, therefore, was due to the fact that subjects experienced both conditions within an experimental session. Inspection of Figure 2 suggests that whereas performing ST prior to R enhanced R performance

to some extent, performing R prior to ST was detrimental to ST performance. It would seem, therefore, that the difference between the R and ST functions in Figure 1 reflects a change in subject strategies brought about by performing the tactile memory task under two disparate conditions. The absence of an interaction between activity and retention interval and the ordering of the curves in Figure 2 favours a strategy-change interpretation over an interpretation phrased in terms of memory trace dynamics. A first guess is that the strategy change was a shift in subject's criteria for what was an acceptable reproduction of the locus of stimulation.

Experiment III

The results of Experiment II clearly showed forgetting of tactile material under both R and ST conditions. To what can this forgetting be attributed? The third experiment was conducted to assess whether the forgetting obtained in the present STM situation was due to proactive interference (PI). In verbal STM experiments very little, if any, forgetting is recorded on the first test of a series of STM tests (e.g. Turvey, Brick and Osborn, 1970); however, forgetting is obtained on subsequent tests as a function of the number of prior tests and the similarity of the prior test material (see Wickens, 1970). If the retention of tactile material in Experiments I and II were affected by PI, then a series of STM test with retention interval held constant should yield increasing error in the reproduction of the locus of stimulation.

Procedure

Sixty-four females served as subjects in Experiment III. In contrast to Experiments I and II, a between-subject design was used to eliminate the problem of a differential improvement with practice as found in those two previous experiments and to control for possible residual PI effects across conditions. Stimuli were delivered to each subject's arm by means of the apparatus described above.

Each subject, whether in R or ST condition, was given eight trials which was judged to be sufficient since the greatest amount of PI generally accumulates over the first three trials in STM experiments (see Melton, 1963; Loess, 1964). Normally, with verbal material, the recall decrement over these trials reaches approximately 80%. Locus range was 3.50 in with an average distance between loci of 0.50 in. A 10 s retention interval was used for both R and ST. The eight stimulus locations were counterbalanced across subjects to avoid order effects.

Instructions were presented by means of a Kodak carousel projector scheduled by a tape timer. Each subject was seated at a right angle to the table supporting the stimulation apparatus. During the entire experiment, the subject was required to fixate on the eye-level, slide-projection area on a wall 4 ft in front of her. Each trial followed the same procedure. An asterisk (*) of 2-s duration was the "Ready" signal followed by a 3-s blank, red slide informing the subject that she would be receiving the stimulus at that time. The tactile stimulus was followed in the R condition, by five blank slides, one every 2 s, and in the ST condition, by five 2-s duration slides each containing one two-digit number; the subject's task was to say the two digits aloud, add them, and classify the sum as "odd" or "even". A question mark lasting 15 s served as the "recall" cue. Finally, the intertrial interval, marked by a blank slide, was of 10-s duration during which time the experimenter recorded the subject's response and adjusted the pointer for the next trial.

TABLE I

Mean deviations in inches

Conditions	STM test							
	1	2	3	4	5	6	7	8
ST	0·540	0·675	0·660	0·600	0·560	0·630	0·440	0·555
R	0·560	0·610	0·460	0·700	0·640	0·570	0·520	0·530

Results and Discussion

The data of Experiment III are given in Table I. The data did not suggest any systematic increase in error as a function of STM tests, i.e. there did not appear to be any systematic PI effect of prior tactile stimuli on the retention of subsequent tactile stimuli in the present experimental situation. At best the data implied a decline in performance over the first few tests followed by an improvement in performance over the later tests. In PI experiments of this sort with verbal material, it is not uncommon for performance to improve after an initial decrement, the effect is usually attributed to practice in the memory task overcoming the interfering effects of prior material. A t-test, therefore, was conducted on the difference between Tests 1 and 2 averaged across conditions since this is commonly the region in which the PI effect is most severe. The test revealed no significant decrement from Test 1 performance to Test 2 performance, $t = 1·38$, $df = 63$, $P > 0·05$.

In addition, an analysis of variance was conducted for conditions (ST and R) and STM tests. Neither conditions, nor tests, nor the interaction reached an F of one. In sum, the absence of an effect of tests strongly suggests that the forgetting of a single tactile input, as evidenced in the present series of experiments cannot be accounted for by PI. This must not be taken as affirmation that PI does not occur in tactile STM. Rather, it suggests that PI contributed little, if anything, to forgetting in the particular STM paradigm of the present investigation. The presence of PI might be observed under different conditions of presentation, e.g. delivering in close succession several tactile stimuli to different loci and then asking after a brief delay for recall of the most recent stimulus.

In view of the results of Experiment II, no difference was expected between R and ST since in Experiment III a subject experienced only one of the two conditions. Summing across tests and taking the mean error gave virtually the same score of R and ST, R = 0·57 in, ST = 0·58 in.

A correlation was performed on mean recall scores and mean ST scores. Since this had not been intended from the outset of the experiment the ST scores were not as sensitive as would have been preferred. What the experimenter had recorded during the retention interval was the number of combined sum and odd-even reports made by the subject. Thus, the ST scores were measures of frequency without being measures of accuracy. A similar failing prevented the computation of a correlation on the ST data of Experiment II; there, the task of adding and

classifying the sum of five pairs of digits did not lend itself to quantification. Indeed, in a number of the conditions subjects did not even complete a single list. However, it was felt that the data of Experiment III could be used to provide another assessment of the relation between retaining tactile stimulation and performing concurrently an arithmetic task. The correlation between the number of ST reports and accuracy of recall was $r = -0.44$ which was significant ($P < 0.05$). The implication therefore is that there was some sharing of a common capacity by the two tasks and that subjects traded-off that capacity between the two. Against this, however, stands the fact that in the present experiment recall accuracy in the ST condition did not differ significantly from that in the R condition.

Conclusions

In sum, the retention of the locus of a tactile stimulus delivered to the upper-side of the arm decreased as a function of retention interval, the forgetting asymptoting after approximately 5 s. Although degree of forgetting was affected by retention interval activity, rate of forgetting was not so affected. Some evidence pointed to a partitioning of processing capacity between the retention task and the interpolated activity but overall it was unclear as to how performing ST during the retention interval influenced a subject's memory of the locus of stimulation. Since a significant effect of ST on recall accuracy was only evident subsequent to a subject having already performed under R instructions, the ST effect could not be attributed solely to a reduction in rehearsal capacity (cf. Gilson and Baddeley, 1969). It was suggested that the ST effect of the present experiments may have arisen, in part, from a change in subjects' criterion for what was an acceptable response.

In addition, the forgetting observed in both R and ST conditions could not be accounted for by PI effects, and in the absence of data to the contrary a decay interpretation of the forgetting of discrete tactile stimuli appears acceptable. And finally, it should be noted that in the present experiments many subjects reported using gross descriptions of stimulus location such as high, middle or low area of the arm, but such indicants could hardly have accounted for the degree of recall accuracy observed.

References

BROWN, J. (1958). Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology*, **10**, 12–21.

DILLON, R. F. and REID, L. S. (1969). Short-term memory as a function of information processing during the retention interval. *Journal of Experimental Psychology*, **81**, 261–9.

GILSON, E. Q. and BADDELEY, A. D. (1969). Tactile short-term memory. *Quarterly Journal of Experimental Psychology*, **21**, 180–4.

LOESS, H. (1964). Proactive inhibition in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, **3**, 362–8.

MELTON, A. W. (1963). Implications of short-term memory for a general theory of memory. *Journal of Verbal Learning and Verbal Behavior*, **2**, 1–22.

PEPPER, R. L. and HERMAN, L. M. (1970).   Decay and interference effects in the short-term retention of a discrete motor act. *Journal of Experimental Psychology Mono-garphs*, **83**, No. 2, Part 2.

PETERSON, L. R. and PETERSON, M. J. (1959).   Short-term retention of individual verbal items. *Journal of Experimental Psychology*, **58**, 193–8.

TURVEY, M. T., BRICK, P. and OSBORN, J. (1970).   Proactive interference in short-term memory as a function of prior-item retention interval.   *Quarterly Journal of Experimental Psychology*, **22**, 142–7.

WEINSTEIN, S. and SERSEN, E. A. (1961).   Tactual sensitivity as a function of handedness and laterality.   *Journal of Comparative and Physiological Psychology*, **54**, 665–9.

WICKENS, D. D. (1970).   Encoding categories of words: An empirical approach to meaning. *Psychological Review*, **77**, 1–15.

EVIDENCE FOR AN INTEGRATION
THEORY OF VISUAL MASKING

MAX COLTHEART† AND BARBARA ARTHUR

*Department of Psychology, University of Waterloo,
Ontario, Canada*

When letters are superimposed upon a pattern of black and white squares, they are easier to identify when the pattern is regular than when it is random. If backward masking consists of the superimposition of a masking pattern upon the decaying visual trace of a target display, a regular pattern should be less effective as a backward mask than a random pattern. This was found to be so for both multiple-letter and single-letter displays. This result is predicted by an integration theory of visual masking but not by an interruption theory.

Introduction

The identification of a visual stimulus can be impaired by the presentation of a second visual stimulus before or after the target stimulus. Three different forms of this effect, visual masking, can be distinguished; masking by a flash of light, masking by neighbouring contours ("metacontrast") and masking by a patterned stimulus occupying the same retinal areas as the target stimulus. Our paper is concerned with the latter form of masking.

This form of masking (Kinsbourne and Warrington, 1962a, b; Sperling, 1963) is at present of considerable theoretical importance for experiments dealing with visual information-processing. Since a visual representation of a visual display persists for some time after display offset (Sperling, 1960), the time for which such material is visually available is not normally under the control of the experimenter. Such control is often desirable; for example, it is required if one wishes to discover the relationship between the amount of time a visual display is available and the amount of material processed from the display. Sperling (1964, 1971) has argued that the experimenter can gain control over the effective exposure duration of a visual display by arranging for the offset of the display to coincide with the onset of a patterned masking field, which "is intended to obliterate any visual persistence of the (display)" (Sperling, 1964), and which produces "processing cut-off of the first stimulus by the onset of the second stimulus" (Sperling, 1971).

Not everyone agrees that a backward masking pattern does terminate the visual representation of a previously-presented display. Neisser (1967), for example, has denied this. He contended that the visual representation of a display may continue

---

to be processed after the onset of a pattern mask, the effect of the mask being merely to degrade, rather than to abolish, the visual representation.

These different conceptions of the mode of operation of a patterned mask have been discussed by Kahneman (1968), who distinguishes two competing theories; the interruption theory and the integration theory. According to the interruption theory, the identification of the target stimulus takes time, and if the onset of the mask occurs before the point in time at which the target stimulus is identified, identification will be prevented because the target stimulus will be replaced by the mask. Sperling (1971) has referred to this effect as "processing cut-off".

The integration theory, on the other hand, assumes that the masking stimulus summates with, rather than replaces, the target stimulus. For example, Kinsbourne and Warrington (1962b) maintain that, when the time between the onset of the target and the onset of the masking stimulus is sufficiently brief, "the two stimuli are treated as 'simultaneous' and are recorded as a composite image, without temporal differentiation. When this composite image includes a random pattern ... it may prove too complex for retrospective analysis" (Kinsbourne and Warrington, 1962b, p. 245). Similar views have been advanced by Eriksen (1966), initially with reference to masking by a light flash, but subsequently with reference to integration of two patterned stimuli presented sequentially (Eriksen and Collins, 1967).

It is not easy to devise experiments which might enable a decision to be made between these two competing accounts of the nature of backward masking, although Liss (1968) and Spencer and Shuntich (1970) have attempted to do so; their work is discussed below. However, Kinsbourne and Warrington's elegant description of the process by which they consider masking to occur suggests a possible approach to the problem. What is required is a situation in which the composite test-stimulus-plus-mask may *not* prove too complex for analysis. Figure 1 illustrates such a situation. In this figure, the word "STOVE" is superimposed upon two patterns, each consisting of 2500 squares, 1250 of the squares being black and the remainder being white in each pattern. In one pattern, the squares regularly alternate from black to white horizontally and vertically. In the other pattern, the blackness or whiteness of a cell is random.

It is obvious that the random pattern obscures the superimposed letters far more than the regular pattern does.† If masking involves attempting to extract the target stimulus from a composite stimulus comprising target plus mask, the phenomenon illustrated in Figure 1 indicates that it would be much easier to perform such a task if the mask is a regular checkerboard than if it is a random checkerboard. Consequently, the integration theory of backward masking would predict that a regular checkerboard would be a less effective mask than a random checkerboard. If the interruption theory is correct, there is no reason to expect any difference in the efficacy of the two masking patterns, provided that they are equivalent in their physical characteristics (e.g. for both of them 50% of the cells are white and 50% black, the contrast of black to white is the same for both masks, and the cell sizes in the two masks are equal).

† The reason why this happens, though obviously of great intrinsic interest, is not relevant to the present paper. What is important for our purposes is the way in which the effect can be used as a tool for studying backward (or forward) masking.

Experiment I

Materials and apparatus

Two checkerboard patterns were constructed by blackening 313 of the 625 squares in a 25 × 25 matrix measuring 9 × 9 in (1 in = 25·4 mm). For one pattern (the regular mask) this was done systematically so that a regular checkerboard was produced. For the other pattern (the random mask) the 313 squares to be blackened were chosen at random. The two matrices were photographed and reduced until the width of one cell was equal to the
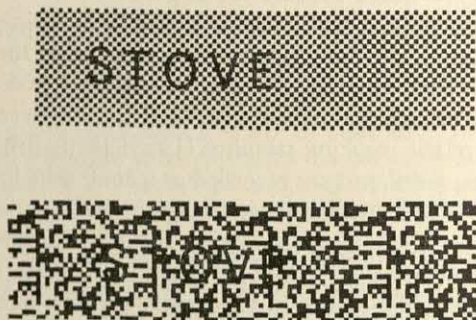
FIGURE 1. An illustration of the differential effects of superimposed regular and random patterns.

width of the capital I in Letraset 287 characters. Rectangular masking patterns were produced by abutting several square patterns in a row. Examples of the regular and random masks are given in Fig. 1. Measurements made by means of a Pritchard Spectra photometer indicated that the two types of mask were equivalent with respect to contrast and average brightness.

Forty-eight 8-letter rows were printed in Letraset 287 on tachistoscope cards. The set of sequences was constructed so that all letters except Q and O were used twice in each of the eight row positions and no letter appeared more than once in a single sequence. Each letter subtended 13' arc vertically. The centres of adjacent letters were separated by 28' arc. The mask measured 0·85° × 2·55°.

A Scientific Prototype Model GB three-field tachistoscope was used for presentation of visual displays. The brightness of each field was 48 ft-L., measured by means of a Pritchard Spectra photometer with a blank white card in the tachistoscope field.

Procedure

After 5 practice trials, 48 main trials were run, each of the 48 letter rows being used once. A fixation point was continuously visible except when a letter row or a mask was being presented. On each trial, after the experimenter's verbal ready signal, a letter row was presented for 100 msec, followed by a 2-s presentation of the masking pattern. The subject was asked to report as many letters from the row as he could. On 24 of the trials, the random mask was used; on the remaining 24, the regular mask. The letter rows which were masked by the random mask for odd-numbered subjects were masked by the regular mask for even-numbered subjects and vice versa. The subjects were 24 undergraduates at the University of Waterloo.

Results

The mean number of letters reported per trial was 4·05 (random mask) and 4·45 (regular mask). A related-groups t-test indicated that this difference was highly
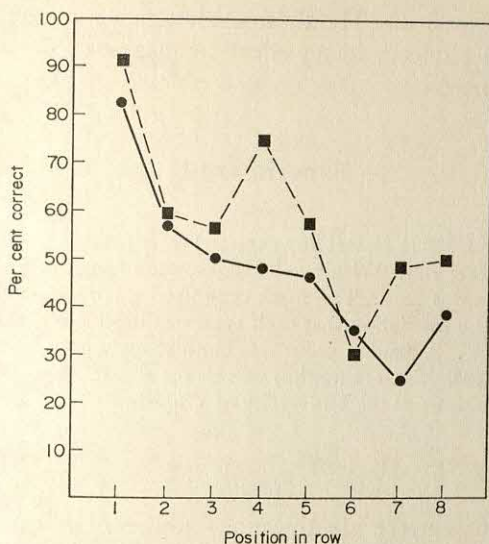
FIGURE 2. Per cent correct report of letters in Experiment I as a function of masking condition and position in row. ■ - - - - ■ Regular (mean = 56·13%); ●——● Random (mean = 50·00%).

significant ($t = 2 \cdot 87$, $df = 23$, $P < 0 \cdot 01$). Figure 2 shows the percentage of letters correctly reported from each position in the row as a function of masking condition. It will be observed that the difference between the two masking conditions is larger in the centre of the row than at the ends. This is an instance of the "selective masking effect" investigated by Merikle, Coltheart and Lowe (1971), and its occurrence in the present experiment is consistent with the tentative explanation which they offered for the effect, namely, that it occurs because the ends of a briefly-exposed letter row are processed before the remainder of the row, and therefore usually before the onset of the mask.

The difference between the random mask and the regular mask was reasonably large, the regular mask, permitting the report of about 10% more material than the random mask, despite the fact that two features of Experiment I might have been expected to reduce the size of the difference between the masking conditions. First, the amount which can be reported from a brief visual presentation may be limited by a non-perceptual factor, loss during report (due, possibly, to output interference). If the number of letters reportable is limited by this kind of factor, differences due to different masks might be reduced or obscured. For example, if six letters can be processed in the regular-mask condition and five in the random-mask condition, but only four can be reported because any letters in excess of this are lost during the report of the first four, the differential masking effects will rarely be manifest in the subject's output.†

Second, in Experiment I both masks were presented for a long period (2 s). It is possible that the relative inefficacy of the regular mask might be further accentuated by presenting the masks for a shorter period.

† This point was brought to our attention by P. M. Merikle.

Consequently, in Experiment II only single letters were presented, and the mask duration was greatly reduced, in an effort to maximize the difference between regular and random masks.

Experiment II

Procedure

The letters E O S T and V (Letraset 287 capitals) were used as targets. For each subject, 10 blocks of 25 trials were run. Within each block, each target letter was used five times, and only one mask type was used. The mask type used was randomized across trial blocks for each subject, with the restriction that each type was used five times. The mask dimensions were 0·85° × 2·55°. A fixation point was continuously present except when target or mask was being presented. Target duration was 20 ms and mask duration was 50 ms. The subjects were undergraduates at the University of Waterloo.

Results

The proportion of correct identifications under each masking conditions is shown, for each subject, in Table I. Individual tests of the difference between the random-mask and regular-mask proportions were carried out. For each subject the regular-mask proportion was significantly greater than the random-mask proportion.

TABLE I

Proportions of correct identifications under two masking conditions in Experiment II

Subject	Regular	Random
1	0·94	0·28
2	0·84	0·34
3	0·87	0·26
4	0·91	0·26

The particular target letters used in this experiment were deliberately chosen. It seemed likely that letters with curved components (O,S) or oblique components (V) would be less strongly masked, with the particular masks we used, than would letters with vertical and horizontal components (E, T), because the masks contain horizontal and vertical contours but no curved or oblique contours. It is evident from Figure 1 that E and T are considerably more obscured by a superimposed random mask than are S, O or V. Unfortunately, we could not pursue this point because of ceiling and floor effects; for all the letters, performance was nearly perfect under the regular-mask condition and nearly at chance level under the random-mask condition.

However, the occurrence of these floor and ceiling effects demonstrated the success of our attempt to amplify the effect obtained in Experiment I; by using a single-letter task and brief mask durations, the superiority of the regular-mask condition was greatly increased.

Discussion

The results of these experiments are predicted by any theory which views backward masking as integration. There seems no reasoning by which an interruption theory could be shown to imply these results. Consequently the results support an integration theory.

Forwarding masking can be viewed in the same way. A visual representation of a masking pattern will persist for some hundreds of milliseconds after mask offset, and if a target array is presented during this period it will be superimposed upon the decaying visual trace of the mask. Once again, the efficacy of any particular masking pattern would then be predictable from the extent to which it obscures the target array when mask and array are presented simultaneously. A major defect of the interruption theory is that it cannot be applied to forward masking, which remains inexplicable, whereas the integration theory explains both backward and forward masking.

In two recent papers, the interruption theory has been advocated in preference to the integration theory. Liss (1968) argued that backward masking interrupts stimulus processing. It is difficult to see why he came to this conclusion, in view of the findings he obtained when using a variety of masking patterns. The patterns he used varied in the amount of masking produced (a result difficult to explain in terms of an interruption theory); moreover, "with the effective mask, the stimulus seems to disappear just as the mask appears, while with the less disruptive mask, the stimulus typically appears 'through' the mask" (Liss, 1968, p. 330). It is thus clear that at least some of the masking patterns used by Liss did *not* interrupt stimulus processing; in fact his statement is notably consistent with a formulation of backward masking as integration of target and mask. His finding that different masking patterns were differentially effective was presumably a consequence of the fact that when the target and mask were integrated it was differentially difficult to extract target from mask.

Spencer and Shuntich (1970) proposed that a backward mask delayed by more than 150 ms has its effect through interruption of processing, whereas a mask presented within 150 ms of the onset of a brief target operates through integration. They deduce this from their data by assuming that "if a pattern mask does not degrade a single-letter display at a given delay, it will not degrade (through integration) a multiple-letter display at the same or longer delays" (Spencer and Shuntich, 1970, p. 201). Their deduction is not a legitimate one because the assumption they make is unwarranted. If sufficient information has been extracted from a display to identify it before the mask is presented, then the integration theory would not predict that the mask would have any effect, because the display has already been translated into some non-visual form. Since it will take longer to process a multiple-letter display than a single-letter display, backward masking will occur at longer mask delays in the former than in the latter case, according to the integration theory.

The finding that with relatively short mask delays, extent of backward masking is positively related to mask energy is, as Spencer and Shuntich state, good evidence for an integration theory. The finding that masks of different energy are equally effective at delays of greater than 150 ms after the onset of a brief target can also be

understood in terms of integration; at this point the post-exposural visual trace of the target has decayed so much that even the low energy mask is sufficient to obscure it entirely when mask and trace are integrated. Spencer and Shuntich concede this point at the end of their paper (p. 203): "The fact that all three masking functions join together when delayed 150 ms could be interpreted either as erasure or summation".

The integration theory would thus seem to give a good account of Spencer and Shuntich's data. The greater backward then forward masking of multiple-letter displays which they found, and which they cite as evidence in favour of interruption theory, actually supports integration theory. Interruption theory cannot explain why there is any forward masking at all, whereas integration theory predicts weaker forward masking (where the mask has partially decayed before the target is presented) than backward masking (where the target is degraded when the mask is presented).

These results would appear to have inconvenient consequences for studies of visual information-processing which attempt to control the period for which a stimulus is visually available by using a backward mask. Many different masking patterns have been used: visual noise (Sperling, 1963), backward ampersands (Mewhort, Merikle and Bryden, 1969); overlapping W's (Spencer and Shuntich, 1970); patterns of oblique lines (Schiller and Wiener, 1963); superimposed O's and Y's (Liss, 1968). The ease with which various letters could be detected with these various patterns superimposed upon them will be different for each type of mask and for each type of letter. The inference that all these patterns would terminate the visual availability of all targets is therefore not permissible. There is not even any guarantee that the masking pattern used by Sperling (1963) would immediately terminate the visual availability of a target. Therefore, as Neisser (1967) has argued before, Sperling's often-quoted conclusion that letters are extracted from a visual display at a rate of one letter 10 ms does not necessarily follow from his data.

References

ERIKSEN, C. W. (1966). Temporal luminance summation effects in backward and forward masking. *Perception and Psychophysics*, **1**, 87–92.

ERIKSEN, C. W. and COLLINS, J. F. (1967). Some temporal characteristics of visual pattern perception. *Journal of Experimental Psychology*, **74**, 476–84.

KAHNEMAN, D. (1968). Method, findings and theory in studies of visual masking. *Psychological Bulletin*, **70**, 404–25.

KINSBOURNE, M. and WARRINGTON, E. K. (1962a). The effect of an aftercoming random pattern on the perception of brief visual stimuli. *Quarterly Journal of Experimental Psychology*, **14**, 223–34.

KINSBOURNE, M. and WARRINGTON, E. K. (1962b). Further studies on the masking of brief visual stimuli by a random pattern. *Quarterly Journal of Experimental Psychology*, **14**, 235–45.

LISS, P. (1968). Does backward masking by visual noise stop stimulus processing? *Perception and Psychophysics*, **4**, 328–30.

MERIKLE, P. M., COLTHEART, M. and LOWE, D. G. (1971). On the selective effect of a patterned masking stimulus. *Canadian Journal of Psychology*, **25**, 264–69.

MEWHORT, D. J. K., MERIKLE, P. M. and BRYDEN, M. P. (1969).   On the transfer from iconic to short-term memory. *Journal of Experimental Psychology*, **81**, 89–94.

NEISSER, U. (1967).   *Cognitive Psychology*.   New York: Appleton-Century-Crofts.

SCHILLER, P. H. and WIENER, M. (1963).   Monoptic and dichoptic visual masking. *Journal of Experimental Psychology*, **66**, 386–93.

SPENCER, T. J. and SHUNTICH, R. (1970).   Evidence for an interruption theory of backward masking. *Journal of Experimental Psychology*, **85**, 198–203.

SPERLING, G. (1960).   The information available in brief visual presentations. *Psychological Monographs*, **74**, No. 11 (Whole No. 498).

SPERLING, G. (1963).   A model for visual memory tasks. *Human Factors*, **5**, 19–31.

SPERLING, G. (1964).   What visual masking can tell us about temporal factors in perception. In *Proceedings of the Seventeenth International Congress of Psychology*, pp. 199–200. Amsterdam: North-Holland.

SPERLING, G. (1971).   Infomation retrieval from two rapidly consecutive stimuli: A new analysis. *Perception and Psychophysics*, **9**, 89–91.

STROUD, J. M. (1956).   The fine structure of psychological time.   In QUASTLER, H. (Ed.), *Information Theory in Psychology*, pp. 174–207.   Glencoe: Free Press.

STIMULUS ENCODING AND ADDITIVITY OF EFFECTS IN A CARD SORTING TASK

BARRY L. LIVELY

Department of Psychology, Bowdoin College

The work of Sternberg (1969) suggests the independence from one another of the times required for encoding a stimulus, matching it against a set of stored representations in memory and responding. A card sorting experiment was performed to test some implications of this position. The variables manipulated were the number of different letters to be sorted into a pile, orientation of the letter on a card, and whether letters encodable according to a shared physical characteristic were to be sorted into the same (within) or different (between) piles. As predicted, orientation of the letter and number of different letters to be sorted/pile affected sorting time only in the between pile condition where the effects were observed to be additive.

Introduction

A method for decomposing reaction time (RT) into a set of independent stages was recently described by Sternberg (1969). Unlike the technique proposed by Donders (1868), Sternberg's "additive factor" method does not require that stages be added or subtracted by changing from one task to another, only that stage duration vary with task requirements. An assumption central to Sternberg's method is that in a factorial experiment two independent variables which affect the same stage will affect its duration in different ways, thus producing an interaction of effects. Two variables affecting different (and additive) stages will produce additive effects. Using a set of variables thought to overlap in various ways in terms of the processes affected, Sternberg (1969) has demonstrated through an appropriate pattern of interactions and additive effects that at least four independent stages can be conceived as intervening between presentation of a stimulus and evocation of a response in an RT task. Consecutively, these stages deal with encoding the presented stimulus, matching the encoded representation with an appropriate item in memory, selecting an appropriate response and organizing the response to be emitted.

The experiment reported here was generated within the framework of this method. In each of several conditions the subject was instructed to sort 32 cards into two piles, with 16 cards/pile. Each card contained one handprinted letter which was either in normal orientation or was rotated 0, 90, 180, or 270 degrees in the plane of the card. A deck containing rotated letters had all four rotations randomly mixed in the deck. One or two letters were to be sorted onto each pile. The letters used were b, d, i, and j, selected for the features common to the pairs, b and d (a loop) and i and j (a hook and dot). Letters to be sorted possessed a

common feature either within piles (b and d vs. i and j, b vs. i) or between piles (b and j vs. d and i, b vs. d). In the within pile condition stimuli can be encoded with respect to the feature unique to a pile (hook or loop), while in the between pile condition stimulus encoding must reflect higher level characteristics of the letter (perhaps its name).

The implications of the different forms of encoding suggest that only some combinations of the conditions will demonstrate additivity (which will refer here to the lack of a significant statistical interaction). Detailed predictions will be made with reference to a formula which follows from Sternberg's (1969) hypothesis:

$$RT = t_{encode} + nt_{match} + t_{respond}$$

where n is a function of the number of features (or items) unique to a pile and $t_{respond}$ is the combined times for selecting and initiating a response. No differential predictions will be made concerning post-match stages; it is assumed that the time required for selecting and initiating a response is the same in all conditions. The value of n is assumed to increase with the number of items/pile in the between pile condition; here it will be assumed that n equals the number of items/pile. In the within pile condition the same feature discriminates piles as readily if there are two items to be sorted as when only one item is to be sorted/pile; hence, n will be assumed to be 1 in this condition independent of the number of items/pile. As an example, in the between pile condition, n is assumed to be 1 when the discrimination is b vs. d and 2 when the discrimination is b and j vs. d and i. In the within pile condition, n is assumed to be 1 for both the b vs. j and b and d vs. i and j discriminations. Encoding and response times are assumed to be independent of n.

As long as n varies, a plot of RT $\times n$ would have the sum of encoding and response times as an intercept and a slope equal to the average amount of time to match one feature (or item) in memory. (Whether RT varies linearly with n or $\log_2 n$ is immaterial to the present argument; for simplicity it is assumed to vary with n.) If n does not vary, the intercept represents the sum of all three components.

The effect of rotation is also predicted to vary with the within vs. between pile distinction. In the within pile condition the discriminating feature (hook or loop) is invariant over rotational transformation; consequently, it should require no more time to encode rotated than normally oriented stimuli in this condition. In the between pile condition a more complex encoding is required and the intercept should reflect the greater difficulty of encoding rotated stimuli. However, after the encoding is made, the times for later operations should be independent of the normal vs. rotated distinction. This would be revealed in equal increases in sorting time for normal and rotated stimuli in the between condition as the number of items to be sorted/pile increased.

In summary, it is predicted that sorting time for the within pile condition will not vary with either number of items/pile or condition of rotation since the availability of the discriminating feature is not sensitive to these manipulations. In the between pile condition, where an encoding specific to the presented stimulus is required, sorting time is predicted to increase with both number of items/pile and condition of rotation, the effects of these conditions being additive since they are assumed to reflect different stages of information processing.

Method

The 17 subjects (one female) were students at Bowdoin College who participated in the experiment as part of a laboratory requirement.

Decks of 32 cards each were constructed by handprinting in pencil one letter/card. Letters (b, d, i, and j) were printed either in a normal orientation or rotated o, 90, 180 or 270 degrees. In decks containing rotated cards the various rotations appeared equally often. In generating decks under each condition all permissible combinations of letters were used. Within each deck all letters used appeared equally often.

Before the experiment began, the subject was instructed always to hold the deck in the same orientation so that rotated letters could not be physically re-oriented. Also, they were encouraged to be accurate in sorting. At the beginning of a trial the experimenter placed slips of paper in front of the subject at his left and right, each slip identifying the letter(s) to be sorted in that pile. The subject was then told whether the cards to be sorted contained letters in normal or rotated orientation. The subject's sorting time was recorded to the nearest $\frac{1}{8}$ s and the number of errors tabulated. In all, the subject worked through 12 decks, the first four of which were considered practice. In the remaining trials all eight kinds of decks were sorted. The order of kinds of decks was changed from one subject to the next to ensure equal representation of conditions across trials.

Results and Discussion

Mean sorting times for each condition were calculated over subjects independent of whether individuals made errors or not. Reference to Table I suggests that while

TABLE I

Frequency of errors in the various experimental conditions : total possible errors = 544

	Within pile		Between pile	
Orientation	1	2	1	2
Normal	0	0	4	10
Rotated	1	0	7	19

frequency of errors varies with task difficulty, even the highest frequency represents a comparatively low error rate (3·5%). Mean sorting times are given in Figure 1.

It appears from Figure 1 that sorting times for between pile conditions were sensitive to number of items/pile and condition of rotation while sorting time for within pile conditions did not vary with either number of items/pile or condition of rotation. These observations were supported by the results of two analyses of variance (Butler, Kamlet and Monty, 1969). The data for between pile conditions were cast into a within-subject analysis where number of items/pile and condition of rotation were the factors. Both numbers of items/pile ($F = 16\cdot50$, $df = 1, 16$, $P < 0\cdot001$) and condition of rotation ($F = 23\cdot59$, $df = 1, 16$, $P < 0\cdot001$) were significant variables. The lack of interaction between the variables ($F = 1\cdot82$, $df = 1, 16$, n.s.) suggests that the variables are additive. A similar analysis for the within pile conditions showed neither variable ($F < 1$ in both cases) nor their interaction ($F = 4\cdot25$, $df = 1, 16$, $P > 0\cdot05$) to be significant. Informal questioning of the subjects after the experiment revealed that methods of encoding were congruent with the predictions.

In view of the lack of interaction, best fitting straight lines were obtained using
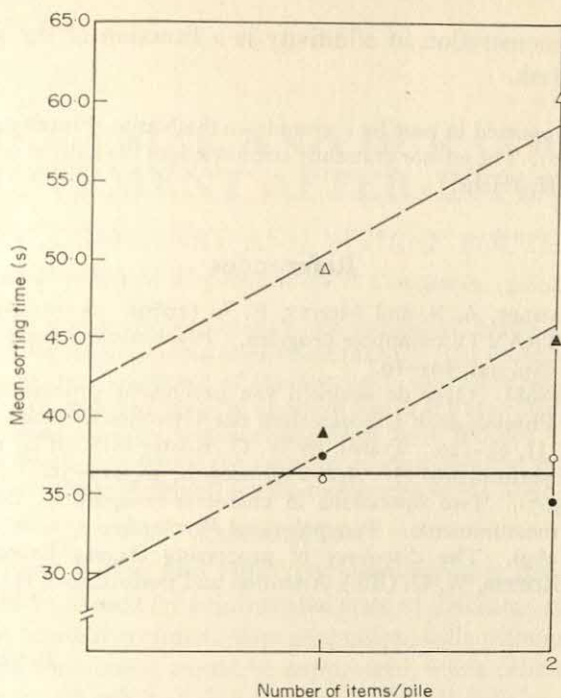
FIGURE 1.   Mean sorting time as a function of number of items/pile for the Normal Within ●——●, Normal Between ▲ - - - ▲, Rotated Within ○——○, and Rotated Between △ - - - △ conditions.

all four data points in the between pile conditions. Since an effect was not found for rotation, only one line was obtained for the within pile data. The slope of this line was set at zero with the intercept representing the mean of the four data points.

According to the hypothesis, the difference in intercept between rotated and normal stimulus functions in the between pile condition reflects the additional amount of encoding time necessary to extract the name from the rotated stimulus. The concomitant lack of interaction between condition of rotation and number of items/pile supports the assumption that matching and response times are independent of encoding time.

The intercept of the within pile function lies above that for the between pile-normal orientation function.   As suggested earlier, the number of components reflected in the two intercepts are very likely different.   The components of the intercept for the between pile-normal orientation functions are assumed to be only the encoding and response times; since the within pile slope is not positive, matching time cannot be estimated and the amount of time to match one feature is thus also reflected in the intercept.   If the matching times were approximately the same in the within and between pile conditions, it would be expected that the intercept would be higher for the between pile function.   The limitations of the present experiment do not permit a more extensive examination of this question.

The results of the present experiment support the Sternberg (1969) hypothesis that encoding and matching times are independent.   Further, the simple point is

made that the demonstration of additivity is a function of the kind of encoding permitted by the task.

References

BUTLER, D. H., KAMLET, A. S. and MONTY, R. A. (1969). A multi-purpose analysis of variance FORTRAN IV computer program. *Psychonomic Monograph Supplement*, **2**, No. 16 (whole No. 32), 301–19.

DONDERS, F. C. (1868). Over de snelheid van psychische processen. Onderzoekingen gedaan in het Physiologisch Laboratorium der Utrechtsche Hoogeschool, 1868–1869, Tweede reeks, II, 92–120. Transl. by W. G. Koster in KOSTER, W. G. (Ed.) (1969). Attention and performance II. *Acta Psychologica*, **30**, 412–31.

STERNBERG, S. (1967). Two operations in character-recognition: Some evidence from reaction-time measurements. *Perception and Psychophysics*, **2**, 45–53.

STERNBERG, S. (1969). The discovery of processing stages: Extensions of Donders' method. In KOSTER, W. G. (Ed.), Attention and performance II. *Acta Psychologica*, **30**, 276–315.

APPARENT SIZE AND DURATION OF A MOVEMENT AFTER-EFFECT

CLAUDE BONNET AND VIVIANE POUTHAS

Laboratoire de Psychologie Expérimentale et Comparée, Associé au CNRS

Four experiments were performed to study the relationship between Emmert's law and the duration of the movement after-effect (MAE). The duration of the MAE increased with increased distance of the test field; this result was shown to be produced by the correlative change in apparent size of the after image. The effect did not occur when cues for distance judgments were reduced. Reducing the duration of the MAE suppressed the variation in its duration at varying distances of the test field. Some implications for the mechanism of the MAE are discussed.

Introduction

Sutherland (1961) and Barlow and Hill (1963) assumed that the movement after-effect (MAE) could be caused by a differential state of discharge in cells responsive to movement in opposite directions: after adaptation, cells responsive to the direction of the adapting movement would be suppressed, while cells responsive to the opposite direction would still be firing spontaneously. It may be assumed that the psychophysically measured duration of the MAE is a measure of the time taken for the firing rate in cells responsive to the direction of the adapting movement to recover its normal value.

Data showing the effect of changes in the characteristics of the adapting movement on the psychophysical duration of the MAE are in good agreement with such a hypothesis (Sutherland, 1961; Richards, 1971). However, if some characteristics of the test field affect the psychophysical duration of the MAE, it would be necessary to explain how this could be integrated into an explanation of the MAE assuming such a physiological process. The experiments reported in this paper illustrate this problem. The increase in the apparent size of an after-image when the distance between observer and projection plane is increased has been referred to as Emmert's law. Our experiments attempt to discover whether a change in the distance between the subject and the test field affects only the apparent size of the MAE or whether it also affects other characteristics of the MAE such as duration. The MAE has many characteristics in common with after-images.

Materials and Method

The four experiments used the same apparatus and general procedure.

Materials

The apparatus was a Dodge-type tachistoscope in which one of the channels was extended with a tunnel allowing the presentation of test fields at 60 cm (adapting distance), 79, 106, 144 and 234 cm. Adapting and test fields had a distal luminance of 10·764 lx (=1 ftc). The test field or projection field consisted of small random black elements against a white

background, made from Letraset paper (Letratone type LT 97). The adapting field was a logarithmic spiral with four black bands rotating at 175 or 120 rpm; the ratio of the black to white parts of the spiral was 1. Except for Experiment III, the diameter of the spiral was 6 cm. It was presented against a black background of 19 × 15 cm.

Procedure

During the adapting period, generally lasting for 30 s, the subject was requested to fixate carefully the centre of the rotating spiral, to avoid ocular movement or blinking, and to maintain clear vision of the field. His main task was to indicate by pressing a key when he ceased to see any MAE. In order to reduce intra-individual variability due to changes in the criterion of response, the subject was instructed to press only when he was absolutely sure he could not detect any further after-movement. This criterion, on the basis of previous experiments, was found to be easy to maintain and to give reliable results.

Experiment I

The first experiment studied the variation in the duration of the MAE as a function of the distance of the projection field.

Materials and method

The adapting field was at a constant distance of 60 cm, and the retinal speed of the adapting movement was 175 rpm. After the 30 s of the adapting period, the subject was presented with the test field located at one of four distances (60, 79, 106, 144 cm) presented in a random order from trial to trial according to the method of constant simuli. After preliminary trials, four experimental trials were run for every distance and every subject. Six subjects took part in the experiment, observing the adapting and test fields in monocular vision with the dominant eye.

Results

The mean durations of MAE are presented in Table I, together with the corresponding standard deviations.

TABLE I

Mean durations and standard deviations of the MAE (s) as a function of the distance of the test field (TF)

TF distances (cm)	Duration of the MAE (s)	
	mean	S.D.
60	10·39	2·57
79	10·68	2·58
106	13·12	3·88
144	14·70	2·60

It may be seen that the duration of the MAE increased linearly with the distance of the test field on which the subject projected the MAE, with a homogeneous inter-individual variability. This linear trend was highly significant, with a Snedecor's F for the remainder for other components of less than 1. In spite of the fact that the adapting movement was the same for all of the four conditions, the duration of the MAE appeared to vary proportionally with the distance of the test field.

The adapting stimulus produced a very clear after-image, which varied in apparent size depending upon the distance of the test-field. Changes in the duration of the MAE could be due either to changes in the apparent size of the after-image or to changes in the apparent distance of the projection screen. This relationship was tested in the next experiment.

Experiment II

Materials and method

Three test distances were used: 79, 144 and 234 cm. The adapting period was reduced to 25 s and the adapting speed to 120 rpm. Ten subjects (students in psychology) were run in two different sessions. During the first session they were run with monocular vision (using their dominant eye), and in the second session with binocular vision.

Each session started with preliminary trials with a test field placed 60 cm away. The mean duration of the MAE was calculated over the last five trials; the number of preliminary trials was varied from subject to subject depending upon the variability of their results.

For each of the two sessions, the main experiment involved five trials for each of the three test distances (79, 144 and 234 cm) randomly presented. For the two first trials at each distance, the duration of the MAE was the only response requested. For the next three trials, after having pressed the key for the duration of the MAE, the subject had to estimate successively the perceived distance of the test field and the perceived size of the after-image.

Estimation of the perceived distance was given in centimetres under the instruction that the distance of the adapting field was 60 cm. Estimation of the perceived size of the after-image (diameter) was also given in centimetres under the instruction that the height of the distal test field was 20 cm. (20 cm was the actual height of the test field; however, the expected apparent diameter of the after-image under Emmert's law was 23·4 cm for a distance of 234 cm. This explains the results reported in Table II. At that distance most of the subjects noticed that the apparent diameter of the after-image was larger than the height of the test screen.) Because variations in the perceived size of the after-image were noticed during the course of one trial, subjects were instructed to estimate the diameter at its largest during the trial.

Results

The results are reported in Table II. The duration of the MAE increased

TABLE II

Mean duration, estimated distance and estimated size of the MAE as a function of the distance of the test field

Distances	79 cm	144 cm	234 cm
MAE durations (s)			
monocular	9·02	10·36	11·53
binocular	10·31	11·44	12·92
App. distances (cm)			
monocular	54·1	108·7	201·8
binocular	51·6	104·3	211·2
App. sizes (cm)			
monocular	7·99	14·56	23·13
binocular	7·89	15·16	24·05

linearly with the distance of the test field ($F = 5.89$, $df = 2, 18$, $P < 0.01$). The effect of distance was the same for both modes of vision as manifested by a Snedecor's F of less than 1.00 for the interaction of vision × distance. However, as usually found (see Holland, 1965), the duration of the MAE for monocular vision was shorter than for binocular ($F = 14.75$, $df = 1, 9$, $P < 0.005$). The same trend appeared in the results of the preliminary trials where the mean duration for monocular vision was 9.13 s and for binocular vision 9.97, but the difference was not significant.

The apparent distance of the test field increased linearly with the physical distance ($F = 181.52$, $df = 2, 18$, $P < 0.0005$). No significant difference ($F < 1$) appeared between monocular and binocular vision. However the slopes of the effect changed with the mode of vision as manifested by a significant interaction between distance and mode of vision ($F = 3.71$, $df = 2, 18$, $P < 0.05$); the binocular vision condition yielded a steeper slope than the monocular.

As predicted by Emmert's law the apparent size of the after-image increased linearly with the distance of the projection screen ($F = 93.29$, $df = 2, 18$, $P < 0.0005$). Apparent size was not affected by the mode of vision ($F < 1$).

The main result of this experiment was that test distance had a similar effect on the three dependent variables. The results conformed to Emmert's law in the strict sense, even for monocular vision.

Using a correlation analysis, an attempt was made to discover closer relationships between duration of the MAE and apparent size or apparent distance. The results were in no way significant, nor even suggestive. The next experiment was carried out to get more information about these relationships.

Experiment III

In the preceeding experiment it was assumed that using monocular vision would reduce constancy by eliminating some of the cues of distance. However it was found that this effect was not strong enough to affect the emergence of Emmert's law nor to interact with the effect of the change of apparent size on the duration of the MAE. In Experiment III, a more drastic method of reducing distance cues was used in order to suppress the effect of distance on the duration of the MAE.

Material and method

The adapting field used was the same as in the preceeding experiments, except that the size of the rotating spiral was reduced to 3 cm diameter. Its speed of rotation was 120 rpm. Three test distances were used 79, 144 and 234 cm. Viewing was monocular and two conditions were used—a full cue condition similar to that used in the preceeding experiments, and a condition of reduced vision. Under the latter condition a black reduction screen containing a circular hole of 3 cm diameter was interposed between the subject and the test field. This screen was placed 60 cm away so that for every distance the visible surfaces of the test field corresponded exactly to the size of the after-image as predicted by Emmert's law. In such a condition no cues of distance from the perspective of the tunnel were available. However, as in the preceeding experiments, the angular size of the "noise" elements changes with distance; increasing the distance also reduces the apparent brightness of the test field and reduces the discriminability of the "noise" elements. For each of two subjects, the experiment was run in four different sessions, alternating from session to session the conditions of vision (full cue or reduced). In every session the three distances were presented in random order six times and the results of the five last trials of each session were analysed. The duration of the MAE was measured.

Results

The results are presented in Figures 1 and 2 separately for each subject. It may be seen that in the full-cue condition the duration of the MAE increased linearly

FIGURE 1. Duration of the MAE as a function of the distance of the projection field under full cue condition (F) and reduced condition (R) for subject CB.
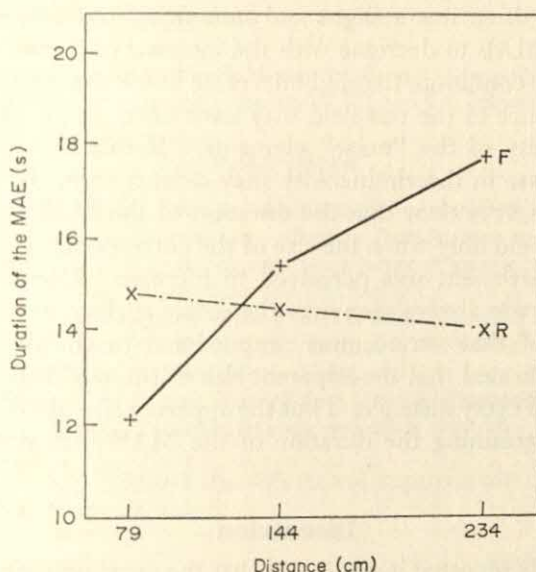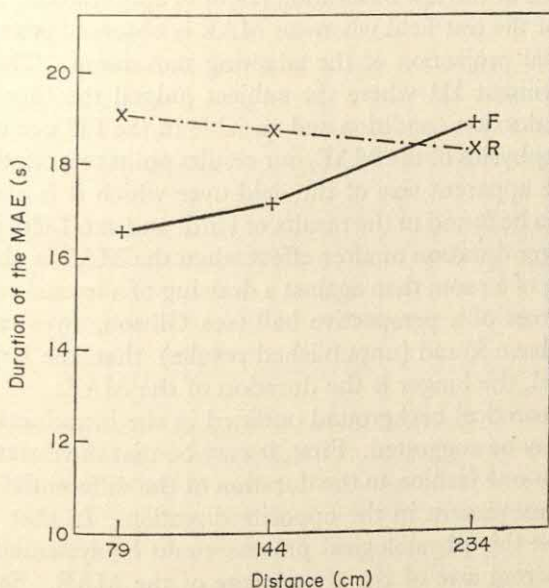

FIGURE 2. Duration of the MAE as a function of the distance of the projection field under full cue condition (F) and reduced condition (R) for subject LH.

with the distance of the test field as found in the previous experiments. The linear regression was significant (for subject 1, $F = 77.61$, $df = 1, 54$, $P < 0.0005$; and for subject 2, $F = 7.50$, $df = 1, 54$, $P < 0.001$) while for both subjects deviation from the linear regression gave a Snedecor's $F < 1$. However, in the reduced condition of vision, there was a slight and non-significant tendency ($F < 1$) for the duration of the MAE to decrease with the increase in distance.

In the reduced cue condition, the slight decrease in the duration of the MAE with the increase in distance of the test field may have been produced by the decrease in the discriminability of the "noise" elements. If this is true, in the full cue condition the decrease in discriminability may detract from the size of the main effect. Nevertheless, it is clear that the duration of the MAE increased with the distance of the test field only when the size of the corresponding after-image of the adapting field of movement was perceived to increase. Under the reduced cue condition the MAE was always seen over the whole surface seen through the hole and the diameter of that surface was proportional to the distance. However subjects' reports indicated that the apparent size of the test field seen through the hole was the same for every distance. Thus the apparent size of the after-image is the critical variable determining the duration of the MAE with distance of the test field.

Discussion

In the experiments reported it was found that the psychophysical duration of the MAE increases with the distance of the test field in conditions where Emmert's law is also valid. The increase in the apparent size of the after-image which is generally related to size constancy could also, in the case of an after-movement, be predicted on the basis of the law of location (Bonnet and Pouthas, 1972). This law states that the size of the test field where an MAE is observed is angularly equal to the size of the retinal projection of the adapting movement. This law remained valid even in Experiment III where the subject judged the three apparent sizes to be equal in the reduction condition and variable in the full cue condition. With regard to the psychophysics of the MAE, our results point to a relationship between its duration and the apparent size of the field over which it is observed. Such a relationship may also be found in the results of Hildt and van Liere (1965) who have demonstrated a longer duration of after-effect when the MAE is observed against a perspective drawing of a room than against a drawing of a frontal wall. Similarly in using different degrees of a perspective hall (see Gibson, 1950) as a form of the Ponzo illusion, we have found (unpublished results) that, the larger the apparent size of the test spiral, the longer is the duration of the MAE.

Following the theoretical background outlined in the introduction, two alternative explanations may be suggested. First, it may be that the duration of the MAE is related in a one-to-one fashion to the duration of the differential state of firing in cells responsive to movement in the opposite direction. In that case, one would have to explain how this physiological process could be systematically altered by changes in the apparent size of the after-image of the MAE. Second, and more probably, the psychophysical duration of the MAE may be directly related to the duration of the assumed physiological process only in the case where the test field

remains constant. This physiological process, even if necessary, is not sufficient to explain the phenomenal appearance of the MAE. In other words, physiological process itself may not give rise to a perception, but to physiological information which would be integrated at a higher level in the visual system with other information in order to generate a perception.

References

BARLOW, H. B. and HILL, R. M. (1963). Evidence for a physiological explanation of the Waterfall phenomenon and figural after-effects. *Nature*, **200**, 1345–7.

BONNET, C. and POUTHAS, V. (1972). Interactions between kinetic and figural information in movement after-effect. *Perception and Psychophysics*. In press.

GIBSON, J. J. (1950). *The Perception of the Visual World*. Boston: Houghton Miflin.

HILDT, M. T. and VAN LIERE, D. E. (1965). Influence of test figure depth cues on the spiral after-effect. *Perceptual and Motor Skills*, **20**, 392.

HOLLAND, H. C. (1965). *The Spiral After-effect*. Oxford: Pergamon Press.

RICHARDS, W. (1971). Motion detectors in man and other animals. *Brain, Behavior and Evolution*, **4**, 161–81.

SUTHERLAND, N. S. (1961). Figural after-effects and apparent size. *Quarterly Journal of Experimental Psychology*, **13**, 222–8.

ACQUISITION, EXTINCTION AND REACQUISITION OF A CONDITIONED RESPONSE IN THE COCKROACH: THE EFFECTS OF OROTIC ACID

J. T. RICK[†], G. W. O. OLIVER[‡] AND G. A. KERKUT

Department of Physiology and Biochemistry, The University of Southampton

Cockroaches can be trained to keep their metathoracic leg out of a saline solution. Isolated metathoracic ganglia learnt faster than headless animals, which in turn learnt faster than intact animals. Extinction took longer in the isolated ganglion than in the other two preparations. Extinction times increased with increasing dose of orotic acid, a precursor of RNA. Orotic acid did not systematically affect the times for acquisition and reacquisition of the learnt response.

Cockroaches will hold a leg in a flexed position for increasing periods of time if the leg is shocked when extended. A preparation first described by Horridge (1962) demonstrated that this withdrawal is a form of avoidance learning, by providing a control animal that was yoked to the learning, or experimental, animal. When the experimental animal extended its leg both animals received a shock. Under these conditions the control animal did not acquire the withdrawal response. The learning can take place in intact animals, headless animals and in isolated segments containing a leg and the ganglion controlling that leg's movements (Aranda and Luco, 1969). The preparation has been used to investigate neurophysiological (Hoyle 1965) and neurochemical (Kerkut, Oliver, Rick and Walker, 1970a, b; Oliver, Taberner, Rick and Kerkut, 1971) changes associated with the avoidance learning.

Here we describe our studies of the response during acquisition, extinction and relearning in intact animals, headless animals and isolated segments. Following a report (Matthies, 1969) that orotic acid, an RNA precursor, can increase the extinction time of a conditioned response in the rat by more than 100%, while having little or no effect on the time needed for acquisition, we studied the effects of orotic acid on acquisition, extinction and reacquisition of the leg withdrawal response in the three types of cockroach preparation.

Method

Adult cockroaches, *Periplaneta americana,* were obtained from a local dealer. Each animal was fixed with elastic bands to a wax block and positioned so that its metathoracic leg was just able to enter a saline bath. A pen-recorder measured the number of shocks received. Once an animal had kept its leg in the saline for 90% of any 3-min period, a stimulator was switched on to give shocks to the leg while contact was maintained. The stimulus was of 60 V and 2 ms duration, at the rate of one shock every second. These conditions continued

---

[†] Present address: Department of Psychology, University of Birmingham, Birmingham, 15.
[‡] Present address: The School of Pharmacy, Portsmouth Polytechnic, Portsmouth.

until the animal received less than four shocks in any successive 3-min interval, at which point the stimulator was switched off. The animal was allowed to extinguish. The criterion for extinction was again maintenance of the leg in the saline for 90% of any 3-min period. This criterion having been reached, the time of reacquisition of the withdrawal response was measured.

Three preparations were studied: intact animals, headless animals and isolated segments. Isolated segments were prepared by cutting the connectives either side of the metathoracic ganglion following ventral incisions. The experiments were set up away from draughts and the temperature maintained at 20 °C since the rate of learning has been shown to be sensitive to small temperature changes (Kerkut et al., 1970b).

All animals were injected, while intact, 24 h before taking part in an experiment. The drug concentrations were made up in Ringer's solution and injected into the haemocoele of the animal. Non-drugged animals had an equivalent volume of Ringer's solution injected. The Ringer's solution used had the formula: 214 mM NaCl, 3·1 mM KCl, 1·8 mM CaCl₂, 1·0 mM TrisHCl; pH 7·2, and was made up in accordance with a modification of Yamasaki and Narahashi (1959). The amounts of orotic acid injected were 0·25 mg, 0·5 mg or 1·0 mg.

Results

The mean times, together with standard errors, for acquisition, extinction and reacquisition of the withdrawal response with each of the three preparations under no-drug conditions are illustrated in Figure 1. It can be seen that as the amount of nervous tissue decreases, so too does the time needed to reach the criterion for acquisition. These differences in the initial learning rates are more clearly observed in Figure 2, which presents average learning curves for the intact animal, the head-
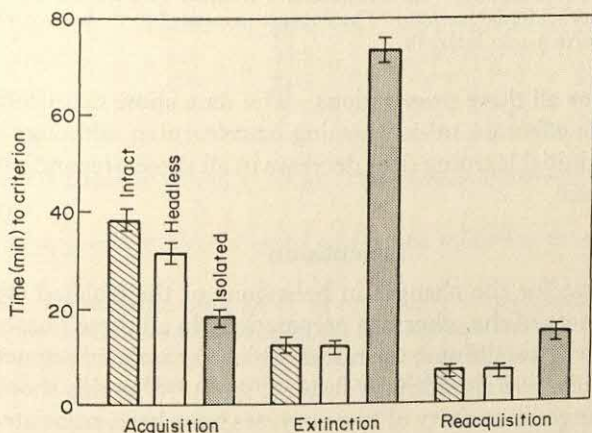
FIGURE 1. Times of acquisition, extinction and reacquisition of a learnt response in three cockroach preparations. Each group consisted of 15 animals. Each bar of the diagram represents a mean and its standard error.

less animal and the isolated segment. Conversely, the time needed for extinction in the isolated segment, Figure 1, is significantly greater than that required by either of the other preparations. Also, it is only in the isolated segment that the relearning time does not differ noticeably from the initial learning time.

Table I compares the effects of the three doses of orotic acid on acquisition, extinction and reacquisition with the no drug conditions. Extinction time increased

FIGURE 2. The curves represent the accumulative number of shocks received by three cockroach preparations during initial learning. The data are presented as the mean number of shocks for 15 animals at successive 3-min intervals.

with dose-level for all three preparations, The data show that orotic did not have similar systematic effect on initial learning or relearning, although at the 1000 μg dose the time for initial learning does decrease in all three preparations, markedly so in the intact animal.

Discussion
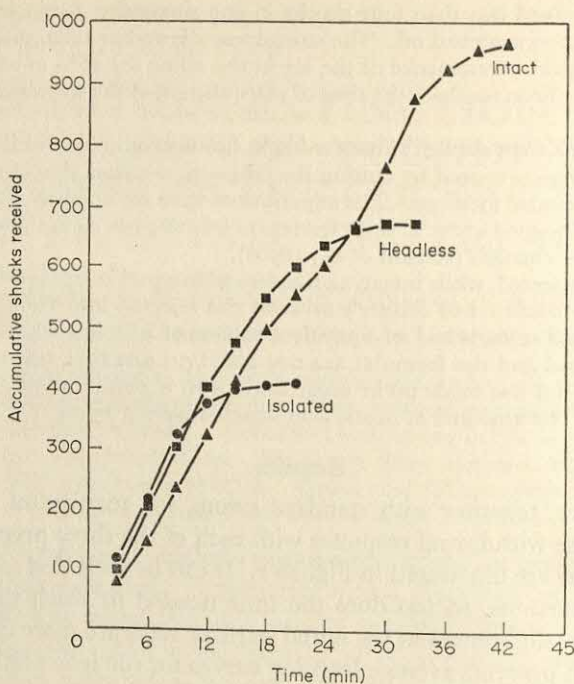
The time course for the changes in behaviour of the isolated segment differed markedly from those of the other two preparations in all three phases of the experiment under no-drug conditions; the rate at which the isolated segment acquired the withdrawal response was significantly faster (Fig. 2). Why this should be the case is not known. Changes in activity of two enzymes have been associated with acquisition of the response in the headless and intact preparations (Oliver et al., 1971). Acetylcholinesterase (AChE) and glutamic acid decarboxylase (GAD) decrease in activity during initial learning in both the ganglion directly associated with the movement of the leg and other ganglia. AChE and GAD are respectively associated with excitatory and inhibitory transmission in cockroach CNS (Pitman and Kerkut, 1970). Such enzymic changes could be due to a side-effect of the training procedure and experiments are in progress to establish whether the enhanced learning rate in the isolated segment is associated with similar, but increased rates of change, in enzymic activity.

Since the time needed for reacquisition of the learnt response in the isolated

segment did not differ significantly from that needed for acquisition (Fig. 1), it may be argued that in this preparation a passive mechanism akin to "forgetting", rather than an active repression of the acquired response, took place during the extinction phase. Such an argument would be consistent with the relatively long period needed for "extinction" by the isolated ganglion. However, orotic acid did have a systematic effect on the rate of behavioural change during the extinction phase in all three preparations (Table I).

TABLE I

Acquisition, extinction and reacquisition of a leg withdrawal response in the cockroach: the effects of orotic acid

Preparation	Orotic acid (μg)	Acquisition Time (min)	Extinction Time (min)	Reacquisition Time (min)	N
Isolated	Control	$18 \cdot 0 \pm 2 \cdot 0$	$73 \cdot 7 \pm 1 \cdot 5$	$15 \cdot 9 \pm 1 \cdot 8$	15
	250	$16 \cdot 5 \pm 2 \cdot 5$	$79 \cdot 7 \pm 9 \cdot 3$	$16 \cdot 2 \pm 2 \cdot 3$	6
	500	$19 \cdot 3 \pm 1 \cdot 8$	$164 \cdot 8 \pm 9 \cdot 4$	$15 \cdot 0 \pm 1 \cdot 7$	6
	1000	$14 \cdot 4 \pm 1 \cdot 3$	$246 \cdot 2 \pm 9 \cdot 1$	$15 \cdot 8 \pm 1 \cdot 2$	6
Headless	Control	$32 \cdot 0 \pm 1 \cdot 4$	$12 \cdot 2 \pm 1 \cdot 1$	$7 \cdot 1 \pm 1 \cdot 9$	15
	250	$30 \cdot 0 \pm 0 \cdot 8$	$48 \cdot 3 \pm 8 \cdot 0$	$11 \cdot 5 \pm 0 \cdot 6$	6
	500	$31 \cdot 9 \pm 0 \cdot 6$	$65 \cdot 0 \pm 1 \cdot 9$	$12 \cdot 0 \pm 0 \cdot 8$	6
	1000	$30 \cdot 5 \pm 0 \cdot 9$	$136 \cdot 8 \pm 4 \cdot 1$	$11 \cdot 3 \pm 0 \cdot 9$	6
Intact	Control	$39 \cdot 0 \pm 2 \cdot 4$	$13 \cdot 7 \pm 2 \cdot 0$	$6 \cdot 6 \pm 1 \cdot 9$	15
	250	$38 \cdot 5 \pm 8 \cdot 8$	$53 \cdot 8 \pm 7 \cdot 0$	$11 \cdot 0 \pm 2 \cdot 1$	6
	500	$32 \cdot 5 \pm 8 \cdot 8$	$61 \cdot 0 \pm 7 \cdot 6$	$9 \cdot 5 \pm 1 \cdot 3$	6
	1000	$28 \cdot 2 \pm 4 \cdot 6$	$114 \cdot 8 \pm 6 \cdot 8$	$7 \cdot 2 \pm 1 \cdot 5$	6

Data are presented in terms of Means $\pm$ s.e.m. The number of animals used is given in the right-hand column.

Note: (a) the dose-dependent effect of orotic acid on the extinction times for each preparation.

(b) the lack of systematic drug effect on acquisition or reacquisition. (The increase in reacquisition times with drug in the headless and intact preparations probably reflect partial forgetting due to the extended extinction periods, rather than a toxic effect, since these increases were not observed in the isolated preparation).

Matthies (1969) showed a similar specificity in the action of orotic acid on the extinction of a conditioned response in the rat. He also demonstrated that the effect did not occur if the incorporation of orotic acid into the pool of uridine phosphates is blocked by 4-aza-uracil. Orotic acid is normally metabolized in nervous tissues into uridine phosphates which are then incorporated into RNA synthesis. Matthies' results suggest that RNA metabolism has some function in the extinction of a conditioned response The possible involvement of RNA metabolism in the process of extinction has previously been demonstrated in two ways. For example, Coleman, Glassman and Wilson (1971) have reported an increased synthesis of RNA in mouse brain during the extinction of an avoidance response: an

increase comparable to that measured during the acquisition of the response. Extinction has also been facilitated in goldfish by the injection of RNA from the brains of extinguished donors (Brand, 1970). This facilitation was not observed when the donor was trained but not extinguished.

The drug data reported here also imply some change in RNA metabolism during extinction. However, our results are anomalous, as are those of Matthies (1969), since orotic acid did not have a similar, systematic effect on the initial learning of the response (Table I), during which enhanced RNA synthesis has been demonstrated (Kerkut et al., 1970a). Preliminary studies of the effects of extinction of the withdrawal response on some neurochemical parameters in the three cockroach preparations has commenced and we hope these will be reported on in due course.

References

ARANDA, L. C. and LUCO, J. V. (1969). Further studies on an electrical correlate to learning. Experiments in an isolated insect ganglion. *Physiology and Behaviour*, 4, 133–7.

BRAND, W. G. (1970). Extinction in goldfish: facilitation by intracranial injection of RNA from brains of extinguished donors. *Science*, **168**, 1234–6.

COLEMAN, M. S., GLASSMAN, E. and WILSON, J. E. (1971). Incorporation of uridine into polysomes of mouse brain during extinction. *Nature*, **229**, 54–5.

HORRIDGE, G. A. (1962). Learning of leg position by headless insects. *Nature*, **193**, 697–8.

HOYLE, G. (1965). Neurophysiological studies on learning in headless insects. In TREHERNE, J. E. and BEAMENT, J. W. L. (Eds), *The Physiology of the Insect Central Nervous System*. New York: Academic Press.

KERKUT, G. A., OLIVER, G. W. O., RICK, J. T. and WALKER, R. J. (1970a). Chemical changes during learning in an insect ganglion. *Nature*, **227**, 722–3.

KERKUT, G. A., OLIVER, G. W. O., RICK, J. T. and WALKER, R. J. (1970b). The effects of drugs on learning in a simple preparation. *Comparative and General Pharmacology*, **1**, 437–83.

MATTHIES, H. (1969). The role of RNA-precursor in memory-processes. *Fourth International Congress of Pharmacology* (*Abstracts*). Basel: Schwabe.

OLIVER, G. W. O., TABERNER, P. V., RICK, J. T. and KERKUT, G. A. (1971). Changes in GABA level GAD and ChE activity in the CNS of an insect during learning. *Comparative Biochemistry and Physiology*, 38B, 529–35.

PITMAN, R. M. and KERKUT, G. A. (1970). Comparison of the actions of iontophoretically applied acetylcholine and gamma aminobutyric acid with the EPSP and IPSP in cockroach central neurons. *Comparative and General Pharmacology*, **1**, 221–30.

YAMASAKI, T. and NARAHASHI, T. (1959). The effects of potassium and sodium ions on the resting and action potentials of the cockroach giant axon. *Journal of Insect Physiology*, **3**, 146–58.

LEARNING WHILE TYPING

F. I. MARTON AND G. SANDQVIST

Institute for Educational Research, Göteborg University

Subjects representing two different levels of skill in typewriting participated in an experiment concerning higher mental activities in connection with typewriting. Understanding and learning of the text typed was facilitated simply by the instruction: "Think of what you are typing!". The improvement in learning was accompanied neither by increase in time nor in mean number of errors.

Introduction

Type-writing can be characterized as a transcription skill which "entail(s) the translation of an input sequence of symbols into an output sequence of responses, and . . . a one-one relationship is preserved in the translation". (Shaffer and Hardwick, 1970, p. 424.)

This preservation of the one-one relationship is of crucial importance. It makes possible reading and typing a text without any understanding of its content. The importance of a higher-order processing in typewriting has already been demonstrated (e.g. Hershman and Hillix, 1965; Shaffer and Hardwick, 1968; Thomas and Jones, 1970). Higher-order processing refers, however, in these cases to the utilization of the linguistic structure in the text at the word or phrase level. This does not presuppose that the subject has paid any attention to the relationship between the linguistic units, i.e. that she had made any effort to grasp the "message". Usually reading serves—or ought to serve—to gain insight into the meaning of the text at a higher—or deeper—level than the one represented by the handling of linguistic units.

Reading, then, has a dual function: first, it produces the input for one-one translation; second, it offers the basic data for understanding and learning the structure and content of the text. In typewriting the former is stressed, in learning—in the cognitive sense—the latter. This line of reasoning led us to the question: what is the relationship between the two different functions of reading, when one of them (the transcriptory one) is in use?

In addition to its theoretical importance, the question has great practical interest. As was shown in a previous unpublished work by Sandqvist the most frequent complaint from people with secretaries is that "the girls do not think of what they are typing", "they type anything". In other words, erroneous manuscripts or recorded dictations will result in erroneous copying. This working set can, of course, partly have its roots in the kind of training given in typewriting, where the aim is to type "as fast as possible, with as few errors as possible". Then perhaps whether or not one thinks of what one is typing is merely a question of a set, which can be turned on or off consciously.

As "thinking of what you are typing" is expected to facilitate learning, retention of the text typed will be one of our dependent variables. Thus the questions to be investigated in the present study can be formulated as follows: does instruction to try to understand a text while typing it affect (a) learning of that text; (b) time or errors of typing; and (c) does it interact with the level of competence of the typist?

Method

Subjects

These were 64 girls who—after leaving the "gymnasium"—had taken part in a one-year secretarial course. Half of the group had had 5 h typewriting a week in school throughout the academic year and the other half only 2 h a week. These two sub-groups represented the two different levels of skill in typewriting, used as one source of variation in the experiment.

Text

All the subjects had to type a text taken from a book of typing exercises. The text consisted of 312 words (corresponding to 2100 strokes). It dealt with the situation of immigrants in Sweden.

Retention test

A test was prepared consisting of 9 questions on the above text, of which 5 placed demands on the knowledge of details in the text while 4 stressed the understanding of the context.

Time

The time necessary for task completion was measured by a stop watch to the nearest second. Thus "time" refers to the period between the start and the notification by the subjects that they were finished.

Typing errors

As typing errors the following were counted: wrong letter, missed letter, missed word, letter reversal within words, wrong word or entire row due to improper finger placement on typewriting keys, correct letter typed over wrong letter.

Design

Sixteen subjects from each level of skill were assigned randomly to the experimental group (special instructions—see below) and the others made up the control group. Then we had a 2 × 2 factorial design and no significant differences between the four sub-groups could be detected as regards their previous academic performance (measured by average school-leaving grades).

After instructions had been given the test was handed out face down. At a given signal the subjects turned up the text and began typing. Each subject upon completing the task raised her hand and a research assistant removed both the manuscript and the text. Upon completion of the task by all subjects, the retention test was given. There was no time-limit on the test.

Instructions

Standard instructions ("type as well and as fast as you can") were given to all subjects. Those belonging to the experimental group were also told to think of what they were typing.

Results

A summary of the results of the arithmetic means for the three dependent variables is given in the Table for the four different subgroups.

TABLE I

Arithmetic means for the dependent variables in four subgroups

Level	Condition	Mean number correct answers in retention test	Mean time(s)	Mean number of typing errors	n
Advanced	Experimental	7·3	1066·3	7·3	16
	Control	6·1	978·8	5·9	16
Less advanced	Experimental	7·3	1515·0	6·8	16
	Control	5·4	1447·5	5·7	16

Three different two-way analyses of variance were carried out corresponding to the three dependent variables. The first one of these, concerning retention test scores, revealed clear differences between the experimental and the control group ($F = 23.99$, $df = 1,60$, $P < 0.001$). Neither the level of skill nor the interaction between instruction and level of skill were found to be significant. The number of correct answers was higher in the experimental group than in the control group on every item except one concerning a year mentioned in the text. This question was correctly answered by all subjects except one in each group.

The analysis of variance on time and number of typing errors respectively showed no significant differences between the experimental and the control group. Thus in spite of the positive effect of instructions on learning, no significant detrimental effects could be detected. The level of skill had of course effect on the time variable ($F = 23.97$, $df = 1,60$, $P < 0.001$), but there was no interaction effect in any of the analyses (the F-ratios were in fact less than one in all cases).

As further support for answering the question about the facilitating effect of instruction on learning, the result of an additional question after the retention test can be seen. All the subject were asked whether they remembered anything more from the text. Twenty-six out of 32 subjects in the experimental group and 5 out of 32 in the control group gave identifiable information on the text. That is a highly significant difference ($\chi^2 = 27.59$, $P < 0.001$).

Discussion

The results seem reasonably clear-cut: it is possible to facilitate higher mental processes while typing without negative effects on the conventional criteria: time and mean number of errors. This is, of course, too strong a statement to be accepted without further replications. The length of the work-period should be taken into consideration, for example. Besides what we mean by knowledge and consequently by retention can vary with the kind of questions we include in our test. Thus one has to be somewhat restrictive when generalizing about the results.

Nevertheless typing offers a very interesting field of research because we can study processes which are going on simultaneously on different levels. These processes can vary from those underlining manual skills to—and this is the main point of this study—higher cognitive activities beyond the level of linguistic units.

References

HERSHMAN, R. L. and HILLIX, W. A. (1965).   Data processing in typing: Typing rate as a function of kind of material and amount exposed.   *Human Factors*, **7**, 483–92.

SHAFFER, L. H. and HARDWICK, J. (1968).   Typing performance as a function of text. *Quarterly Journal of Experimental Psychology*, **20**, 360–9.

SHAFFER, L. H. and HARDWICK, J. (1970).   The basis of transcription skill.   *Journal of Experimental Psychology*, **84**, 424–43.

THOMAS, E. A. C. and JONES, R. G. (1970).   A model for subjective grouping in typewriting. *Quarterly Journal of Experimental Psychology*, **22**, 353–67.

THE OVERTRAINING REVERSAL EFFECT IN RATS: A FUNCTION OF TASK DIFFICULTY

CHARLES L. RICHMAN, KAROL KNOBLOCK AND WAYNE COUSSENS†

Department of Psychology, Wake Forest University, North Carolina

Two experiments were conducted with rats. The first study showed that in the T-maze a brightness discrimination learning problem was more difficult than a spatial learning problem, but in a Ross-maze a brightness problem was less difficult than a spatial task. T-maze brightness and Ross-maze spatial tasks were found to be of equal difficulty.

In the second experiment rats were trained either on a brightness or spatial discrimination reversal problem in the Ross-maze. It was found that overtraining facilitated reversal performance in the spatial task but not in the brightness problem. The theoretical implications of these results were discussed.

Introduction

During the 20-year history of the investigation of overtraining and reversal learning, much conflicting evidence has appeared. Reid (1953) found that reversal learning in a brightness discrimination task was facilitated by additional training trials beyond criterion of original learning. This phenomenon, the facilitation of reversal learning by overtraining, has since been called the overtraining reversal effect (ORE). Other investigators have found that overtraining has no effect on trials to a reversal criterion (D'Amato and Jagoda, 1962; Richman, 1970), or that over-training significantly retards reversal learning, i.e. the reverse ORE (Clayton, 1963; Eimas, 1967, Experiment II; Mackintosh, 1965a). Because of these variant findings, the factors affecting the ORE have recently come under careful investigation and appear to be, when closely scrutinized, more specific than was first supposed.

Lovejoy (1966) has proposed that employing a difficult discrimination task is a necessary condition for the occurrence of the ORE. Lovejoy assumes that the ORE depends on non-overtrained subjects trained on a difficult discrimination task ceasing to attend to the relevant cue during reversal training. But, if the relevant cue is one that an animal automatically attends to, then the non-overtrained subject will continue to attend to the relevant dimension throughout reversal. Because rats trained in T- or Y-mazes generally learn spatial problems in less trials than non-spatial ones it has been assumed that spatial cues in these mazes are attended to auto-matically with the consequence that only two experiments (Capaldi, 1963; Bruner, Mandler, O'Dowd and Wallach, 1958) have reported the ORE employing a simple

† Present address: Department of Psychology, University of Mississippi, Mississippi.

spatial problem. The present report will not concern itself with these two aberrant findings since an adequate explanation has been reported elsewhere (Richman and Coussens, 1970).

One study, related to the issue in question, was performed by Eimas (1967), who found a tendency toward the spatial ORE by using an unconventional discrimination apparatus, a Ross-maze (after Ross, 1962). He also found that rats reached the initial learning criterion in less trials when trained on brightness rather than spatial problems. These results appear consistent with Lovejoy's "difficult discrimination model" that predicts the ORE's occurrence only when employing difficult discrimination tasks.

It is the primary purpose of the present investigation to conduct a more complete test of the "difficulty of discrimination" model by first comparing the relative ease of spatial and brightness discrimination learning in a conventional T-maze to spatial and brightness learning in a Ross-maze. Recent reports by Waller (1968) and Richman, Gardner, Montgomery and Benewicz (1970) suggest that brightness learning rates in a T-maze is equivalent in difficulty to that of spatial learning rates in a Ross-maze and more difficult than a Ross-maze brightness task. If these findings were empirically demonstrated then according to the Lovejoy model one would predict the ORE in the former but not the latter Ross-maze task.

Experiment I

This study compared two types of maze (Ross and T) and two relevant dimensions (spatial and brightness). The primary purposes were: (a) to test the hypotheses that in the Ross-maze spatial learning is more difficult than brightness learning and that in the T-maze brightness learning is more difficult than spatial; (b) to see if the Ross-maze spatial problem is equivalent in difficulty to the T-maze brightness one.

Method

Sixty-four experimentally naive, male albino rats of the Sprague-Dawley strain were the subjects. They were 55–60 days old at the beginning of the study and were purchased from Holtzman Company, Madison, Wisconsin.

Three major pieces of equipment were used in the study, a Ross-maze and two standard T-mazes.

The single unit T-mazes were constructed from 1·27 cm plywood, with no roof and hand-operated guillotine doors placed between the start box and the start stems and 32·6 cm on either side of the choice point in the goal alleys. The maze dimensions were as follows: alleys 12·7 cm wide and 35·6 cm high, starting stems 45·7 cm long, and goal alleys 61·0 cm long, with goal boxes 21·6 cm long. The T-mazes were placed side by side, making the start box positioning very similar to that found in a Ross-maze. The north goal alley of one T-maze and the south alley of the other were painted black on both sides from the edge of the start stem to the point where the guillotine door came down to close off the goal box. The opposite goal alleys were painted white in the same manner.

The Ross-maze was also made of 1·27 cm plywood. Its dimensions are illustrated in Fig. 1. The maze forms two sets of parallel double runways (designated runway set 1 and 2). The start section, facing north-south, and alleys were 45·7 cm deep. All alleys and the start section were 8·9 cm wide, except where the start section narrowed at the entrance to the choice point area, this being 7·6 cm wide. The animal compartment was mounted so that
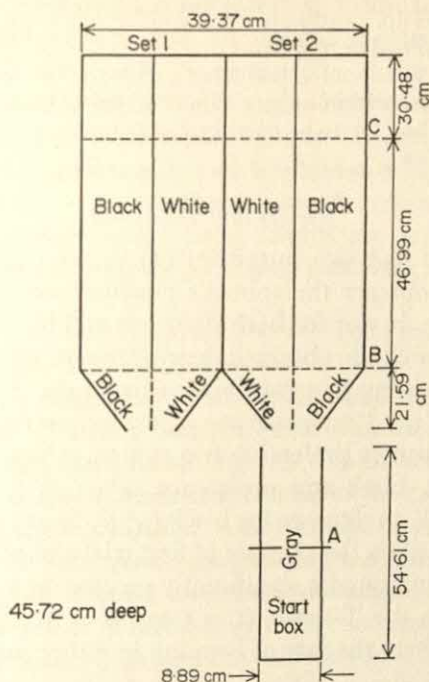
FIGURE 1. Ground plan of the Ross-maze apparatus.

either choice situation could be placed in front of the start section prior to a trial. The start section was painted dull grey. As indicated in the diagram, the alleys at the extreme left and right were painted black and the two middle alleys were painted white. Hand-operated guillotine doors were placed at points A, B, and C (see Fig. 1) permitting (A) restraint in the start box, (B) prevention of retracing, and (C) restraint in the goal box.

The mazes were placed on the floor in a windowless experimental room which measured 2½ by 6½ m. The illumination was by fluorescent lamps which illuminated the room evenly, and each maze was positioned directly under a fluorescent fixture. The room was kept at a constant temperature and humidity. Food reinforcement in both mazes throughout the experiment consisted of 3–100 mg Noyes food pellets placed in a bottle cap located 5 cm from the end of the goal alley.

All animals were maintained on *ad lib* water in individual cages throughout the experiment and were maintained on ad lib food (Purina Laboratory Chow) for three days following their arrival in the laboratory. On the fourth day, the animals were weighed. Daily weighing continued throughout the experiment. The animals were reduced to about 85% of their *ad lib* body weight (±7%). Each rat was given four 100 mg Noyes pellets in his home cage 9–10 h before running (any uneaten lab chow was removed from the cage at this time); this was an attempt to equate for food ingestion time prior to each day's run. This procedure also allowed the animals to become acquainted with the Noyes pellets as a food source. On preliminary training Days 4–9, the animals were handled (picked up 25 times per day). During Days 4 and 5 the animals were given 5 rewarded trials per day in the maze in which they were not to receive discrimination training. On the 6th through 9th day of preliminary training, the animals were given 5 free trials per day in the type of maze in which they were to be trained (each goal box contained 5–100 mg pellets). Position and brightness preferences were noted.

The experimental design consisted of two types of mazes (standard T-maze—T and a Ross-maze—R) and two relevant stimulus dimensions (Spatial—S, and brightness—B),

with 16 animals in each group.   Subjects learned either a spatial (left-right turn) or a visual (black-white turn) discrimination problem.

All animals were given 8 non-correction trials per day.   The learning criterion was set at 7 out of 8 correct responses within a day's run.   Following an incorrect response, subjects were detained in the end box for 10 s.   Start box placements were determined by a Fellows (1967) randomization series.

Results and Discussion

Preliminary training had two purposes: (1) to acclimatize the animals to the apparatus and (2) to observe the animal's position and/or brightness preference behavior.   Preference behavior for both right-left and black-white was determined by a χ^2 test that compared the observed distribution of total responses for the four days of preliminary training to a distribution based on chance expectancy (50%). In neither the T-maze nor Ross-maze was a position preference found ($\chi^2 < 1\cdot00$). However, a black brightness preference was evident; the animals trained in the T-maze showed a 61·7% black arm preference ($\chi^2 = 7\cdot04$, $df = 1$, $P < 0\cdot01$) and Ross-maze rats a 90·6% preference for black ($\chi^2 = 84\cdot50$, $df = 1$, $P < 0\cdot001$). A *post hoc* *t*-test performed on the number of first trial choices summed over the four preliminary test days indicated a significantly greater preference for the black arm in the Ross-maze than the T-maze, ($t = 5\cdot03$, $df = 31$, $P < 0\cdot001$). Preference for black did not facilitate the rate of learning in either maze (P's $> 0\cdot10$).
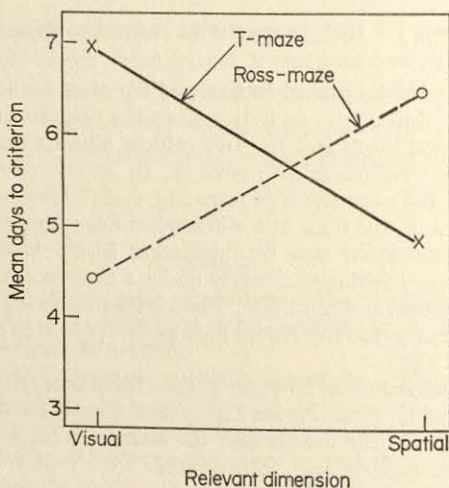

FIGURE 2.   Mean number of days to initial learning criterion as a function of the maze and relevant dimension.

Figure 2 graphically illustrates the mean days to learning criterion. The figure suggests an interaction without main effects.   Indeed, employing a 2 × 2 analysis of variance on the number of days to learning criterion neither statistically significant relevant dimension nor maze main effects were found ($F < 1\cdot0$, and $F = 1\cdot78$, $df = 1,60$, $P > 0\cdot10$, respectively).   A relevant dimension × maze interaction ($F = 10\cdot00$, $df = 1,60$, $P < 0\cdot005$) was evident, however.   A subsequent Duncan Multiple Range Test was performed on data derived from this interaction indicated that the

Ross-maze brightness discrimination problem was easier to master than the Ross-maze spatial problem ($P < 0.01$); and, the Ross-maze brightness task was easier than the T-maze brightness one ($P < 0.005$). The T-maze spatial discrimination problem tended to be easier than the T-maze brightness problem ($P < 0.10$). These findings confirm the *a priori* hypothesis that T-maze spatial tasks tend to be easier to master than brightness ones; and brightness tasks are easier than spatial problems in the Ross-maze. The former result replicated our earlier study (Richman *et al.*, 1970) whereas the latter results are similar to Waller (1968). Finally, the learning rates of Group R-S and T-B were found to be equivalent ($P > 0.20$).

Experiment II

Rat reversal learning studies have shown that the ORE is more consistently demonstrable in difficult T- and Y-maze brightness discrimination tasks than easy spatial learning problems (Mackintosh, 1965*b*). These findings suggest that the ORE may be related to some unique characteristic of the rat's visual system and/or to the relative difficulty of the task used. Unfortunately, previous research has not differentiated between the two hypotheses, leaving the question unresolved. The present experiment was an attempt to test the efficacy of the two hypothesis by using an apparatus that results in faster brightness than spatial learning and makes the latter problem functionally equivalent in difficulty to a T-maze brightness task. Experiment I demonstrated the Ross-maze's adequacy in meeting these requirements.

Method

The subjects were 32 male Sprague-Dawley albino rats, approximately 90 days old. The animals were purchased from Holtzman Company, Madison, Wisconsin.

The general procedures used in Experiment I were employed in the present study with the following exceptions: (1) animals were maintained throughout the experiment on a 23-h food deprivation schedule. This schedule began three days prior to discrimination training; (2) the design consisted of two types of Ross-maze problems (spatial cues relevant with brightness cues irrelevant and brightness cues relevant with spatial cues irrelevant) and two levels of training (criterion and overtraining). Criterion subjects were initially trained to a criterion of 7/8 correct responses in a single day's run; overtrained subjects were trained to criterion and then given eight additional days of training; (3) upon reaching their respective initial learning criteria reversal learning ensued. That is, if an animal was rewarded for turning left during the initial phase he was rewarded for turning right during reversal phase; (4) reversal learning criterion was set at 7/8 correct responses within a day's run.

Results and Discussion

A simple one-way analysis of variance performed on the days to an initial criteron of 7/8 correct responses replicated the results found in Experiment I, faster visual than spatial learning in the Ross-maze ($F = 9.07$, $df = 1,30$, $P < 0.01$). Table I shows the mean days to initial and reversal learning criterion and the mean number of perseverative errors.

The primary purpose of Experiment II was to assess independently brightness and spatial reversal learning rates as a function of training. Therefore, two pre-planned analyses were conducted, one on the brightness and the other on the spatial

TABLE I

Mean days and errors to initial and reversal learning criterion and mean perseverative errors

Groups	Initial learning		Mean number	Reversal learning	
	Mean days	Mean errors	Perseverative errors	Mean days	Mean errors
Brightness criterion	3·50	8·12	3·38	5·75	19·75
Brightness overtrained	3·75	7·75	5·13	4·75	19·38
Spatial criterion	5·25	15·88	3·88	8·50	35·00
Spatial overtrained	6·12	21·38	7·00	5·75	25·25

groups' days to reversal learning criterion. The results of these analyses showed that extended brightness training did not affect reversal learning rates; whereas, extended spatial learning did succeed in facilitating reversal performance ($F = 1\cdot14$, $df = 1,14$, $P > 0\cdot20$; and $F = 8\cdot18$, $df = 1,14$, $P < 0\cdot025$, respectively). These results, a spatial but not a brightness ORE, offer strong support for the "difficulty-of-discrimination" model (Lovejoy, 1966). The latter result is also consistent with Eimas' (Experiment III, 1967) earlier report of a tendency toward the ORE in a Ross-maze spatial learning task.

While the results of the present study are reasonably straight-forward, some qualifications need to be added. Recently, Mackintosh (1969) conducted a well-thought-out series of experiments designed to investigate the factors affecting the ORE. The apparatus Mackintosh (1969) used was a Grice-box of similar design to our Ross-maze. Mackintosh (1969) reported, however, that the spatial original learning task was accomplished in fewer trials than the brightness problem and over-training facilitated reversal learning in the latter task, but not the former. Although the present results are contrary to Mackintosh's (1969) in terms of the relative speed of spatial vs brightness discrimination learning, both studies support the sugestion that the ORE is, at least, partially due to the use of difficult discrimination tasks.

One possible source for this discrepancy among the data lies with the possibility that the mazes used by Mackintosh (1969) and in the present study were sufficiently different to produce differences in the relative speed of spatial vs brightness learning. The suggestion is that the Grice-box used by Mackintosh (1969) provides an animal with less visual stimulation than our Ross-maze. Although the goal arms in both mazes were painted either black or white, the Grice-box arms were only 30·5 cm long and 15·3 cm high; whereas, the Ross-maze length and height were 76·5 cm and 45·7 cm, respectively. Furthermore, in the Grice-box the fan-shaped choice area was painted neutral gray; whereas, the Ross-maze choice area was painted either black, or white, consistent with its adjacent goal arms. Since the declination of body movement needed to enter a goal arm was the same in both

effects of reward magnitude are on perseverative error rates and not attention. mazes, it is assumed that learning difficulty was primarily affected by the amount of differential brightness stimulation provided by the maze and not due to differential proprioceptive feedback at the choice point. This same result was readily demonstrated in two unpublished T-maze studies recently completed in our laboratory. The results of these studies indicated that brightness learning was facilitated by simply extending the goal arm brightness to the center of the choice point. It is worth noting that the ORE was found here, also, only in the more difficult brightness learning task.

The present results taken together with Mackintosh's (1969) suggest what the necessary conditions are and how they interact in producing the ORE. First, it is assumed that overtraining not only strengthens habit strength to the particular reinforced stimulus but also increases the probability of attending to the relevant stimulus dimension (e.g. brightness). Second, previous research has shown that the probability of initial perseverative errors, at the onset of reversal training, is directly related to the amount of original training. (Mackintosh's [1969] Tables 1 and 2 and his statistical analyses of Experiment 3, show that perseverative errors are *not* affected by task difficulty. A *post hoc* t-test performed on the present data confirm this finding ($t < 1.00$, $df = 14$, $P > 0.10$). And, third, it is evident from Mackintosh's results that trials to criterion in both original and reversal learning are directly related. In the present study, we found that criterial trained animals performing the brightness task took less days to reach the original learning criterion than spatially trained subjects. This finding, more rapid brightness than spatial learning, was also evident during reversal training ($t = 2.27$, $df = 14$, $P < 0.05$).

In summary then, (1) assuming that overtraining strengthens attention to the relative stimulus dimension, (2) accepting the fact that overtraining increases perseverative errors at the onset of reversal learning, and that (3) task difficulty does not affect perseverative errors; and, finally, (4) criterion trained animals' original and reversal learning rates are positively correlated we are able to account for much of the overtraining and reversal learning data. The interpretation being offered here is, that, in simple tasks the positive transfer effects due to overtraining are outweighed by negative, perseverative error, effects. Since the ORE is based upon the relative reversal learning rates of criterion trained and overtrained animals and original and reversal learning rates are correlated, we suggest a "floor effect" prevents the ORE's occurrence when easy tasks are employed. That is, in easy discrimination problems by the time (in terms of trials to reversal criterion) overtrained animals' perseverative errors have been extinguished, the criterion trained animals have begun to approach reversal learning criterion. However, by the time overtrained animals' perseverative errors are extinguished in a difficult task and the positive transfer effects due to overtraining begin to show themselves, criterion trained subjects are still responding to the new positive stimulus at about chance level.

This analysis is also relevant to the finding that regardless of task difficulty, the ORE is only demonstrable when large rewards are used. Assuming that the probability of attending to the relevant dimension develops no faster for animals trained with small rewards than with large rewards, we may hypothesize that the

Inspection of Tables 1 and 4 reported by Mackintosh (1969) suggest that perseverative errors are greater for subjects overtrained with small reward than large reward.

With regard to perseverative errors, it appears that the amount of training and reward interact so that reward magnitude has no effect on perseveration when animals are trained only to criterion, but has a substantial effect when animals are overtrained. Both Frustration Theory and Discrimination Hypothesis predict these results.

References

BRUNER, J. S., MANDLER, J. M., O'DOWD, D. and WALLACH, M. A. (1958).   The role of overlearning and drive level in reversal learning. *Journal of Comparative and Physiological Psychology*, **51**, 607–13.

CAPALDI, E. J. (1963).   Overlearning reversal effect in spatial discrimination task. *Perceptual and Motor Skills*, **16**, 335–6.

CLAYTON, K. (1963).   Reversal performance by rats following overlearning with and without irrelevant stimuli. *Journal of Experimental Psychology*, **66**, 255–9.

D'AMATO, M. R. and JAGODA, H. (1962).   Overlearning and position reversal. *Journal of Experimental Psychology*, **64**, 117–22.

EIMAS, P. D. (1967).   Overtraining and reversal discrimination learning in rats. *Psychological Record*, **17**, 239–48.

FELLOWS, B. J. (1967).   Chance stimulus sequence for discrimination tasks. *Psychological Bulletin*, **67**, 87–92.

LOVEJOY, E. (1966).   Analysis of the overlearning reversal effect. *Psychological Review*, **73**, 87–103.

MACKINTOSH, N. J. (1965*a*).   Overtraining, transfer to proprioceptive control, and position reversal. *Quarterly Journal of Experimental Psychology*, **17**, 26–36.

MACKINTOSH, N. J. (1965*b*).   Selective attention in animal discrimination learning. *Psychological Bulletin*, **64**, 124–50.

MACKINTOSH, N. J. (1969).   Further analysis of the overtraining reversal effect. *Journal of Comparative and Physiological Psychology; Monograph*, **67**, 2, 2, 1–18.

REID, L. S. (1953).   The development of noncontinuity behaviour through continuity learning. *Journal of Experimental Psychology*, **46**, 107–12.

RICHMAN, C. L. (1970).   Effects of experimentally induced irrelevant stimuli on simultaneous discrimination problems. *Journal of Genetic Psychology*, **117**, 69–83.

RICHMAN, C. L. and COUSSENS, W. (1970).   The undertraining reversal effect. *Journal of Experimental Psychology*, **86**, 340–2.

RICHMAN, C. L., GARDNER, J., MONTGOMERY, M. and BENEWICZ, K. (1970).   Effects of body weight loss on position learning and brightness tasks. *Learning and Motivation*, **1**, 218–25.

ROSS, L. E. (1962).   The effect of equal reinforcement of the poitive and negative discriminanda of a learned discrimination. *Journal of Comparative and Physiological Psychology*, **55**, 260–6.

WALLER, G. T. (1968).   Effects of magnitude of reward on spatial and brightness discrimination tasks. *Journal of Comparative and Physiological Psychology*, **66**, 122–7.

THE EFFECTS OF STIMULUS DIMENSIONALITY ON THE RATE OF GAIN OF INFORMATION

ROBERT M. LEVY

Indiana State University

AND GLORIA K. NORTON

Bowling Green State University

Hick's paradigm as extended to an absolute judgment task by Doherty was used to study the effects of dimensionality on the rate of stimulus identification in two experiments. Two undimensional conditions, size and brightness, and two bidimensional conditions, size-brightness redundant and size-brightness non-redundant, were employed. The significant linear components of the regression of choice reaction time on transmitted information for the undimensional and bidimensional stimuli supported and extended Hick's law. The finding of nearly identical slopes for the two undimensional and redundant bidimensional condition regression lines suggests a stage of processing which has a constant capacity in bits/s and which is independent of stimulus dimensionality. An increase in slope for the non-redundant condition is attributed to a difference in response requirements.

Introduction

Multidimensional stimuli are formed by combining, with varying degrees of redundancy, values selected from along two or more dimensions. The degree of redundancy indicates the extent to which one value can be predicted from knowledge of the other at a better than chance level. To illustrate, consider the dimensions S and B with two values along each: S_1, S_2, and B_1, B_2 respectively. The two perfectly redundant sets are (S_1B_1, S_2B_2), and (S_1B_2, S_2B_1). Non-redundancy means that values are combined orthogonally, i.e., $(S_1B_1, S_1B_2, S_2B_1, S_2B_2)$. Knowledge of one value does not aid in predicting the other component value. A "semi-redundant" set is (S_1B_1, S_1B_2, S_2B_2).

Many studies designed to investigate the effects of dimensionality on stimulus identification have used the amount of information transmitted in bits per stimulus $(T(S:R))$ as the dependent measure. The general finding for non-redundant combinations (Pollack, 1953; Pollack and Ficks, 1954; Beebe-Center, Rogers and O'Connell, 1955) and redundant combinations (Eriksen and Hake, 1955; Lockhead, 1966) has been that $T(S:R)$ increased as the number of simultaneously varied dimensions increased.

Hick's (1952) finding of an increase in choice reaction time (CRT) as a function of an increase in $T(S:R)$ suggests that a complete analysis of stimulus identification

must include not only an analysis of the amount of information transmitted per stimulus but also of the amount of information transmitted per unit time, i.e., the rate of information processing in bits/s. The primary purpose of the present study was to determine the effects of the number of dimensions and degree of redundancy on the rate of information processing. A second aim was essentially methodological in nature, namely, to separate out from CRT different processing components.

Nye (1962) and Levy (1965) measured both CRT and T(S : R) in their studies of dimensionality. Nye (1962) found that more information was transmitted more quickly with values combined non-redundantly on five dimensions than on three dimensions (total set sizes of 25 and 24, respectively). Levy (1965), with set size fixed at four, measured speed and accuracy of identification of stimuli varying along single dimensions and their redundant and non-redundant bidimensional combination. He estimated the rate of information transmission from the ratio T(S : R)/CRT and found the highest rate with the non-redundant set, the next highest with the redundant, and the lowest with the single dimension sets.

Estimation of the rate of information transmission from the simple ratio of average T(S : R) to average CRT presents difficulties since the CRT contains components which reflect different processes. Assume that on any trial the CRT contains two components, I and K, where I represents a response to the informational content of the stimulus set and K is a constant which includes the response to other characteristics of the stimulus and, possibly a constant contribution of the organism. The ratio estimate of rate does not allow separation of these components. Consequently, use of the T(S : R)/CRT measure biases the estimate of rate to a degree which is a function of the relative magnitudes of I and K.

A procedure which eliminates the bias and permits separation of various components of CRT is one based upon Hick's (1952) finding of a linear relationship between CRT and T(S : R):

$$CRT = a + b \, T(S:R).$$

By taking the reciprocal of the slope (b), the s/bit measure is converted to an estimate of rate in bits/s. The intercept constant, a, contains the components of CRT which lead to bias in the rates estimated from the simple ratio T(S : R)/CRT.

Doherty (1968) generalized Hick's finding from a discrimination task to an identification task. Easy and hard discriminability stimulus sets of unidimensional stimuli varying in linear extent were used. The results confirmed Hick's finding of a linear relation of CRT on T(S : R) for both conditions. The slopes (rates) were the same for both conditions; differences in discriminability were reflected in the intercepts. It appears from Doherty's results that the analysis of the regression of CRT on transmitted information allows the desired separation of various components of CRT which is fundamental to the understanding of the role of different amounts and kinds of dimensionality in stimulus identification.

Experiment I

In Experiment I Hick's paradigm was used to study information processing for unidimensional and redundant and non-redundant bidimensional stimulus sets.

Method

Subjects

One male and one female graduate student (both age 22) from Indiana State University served as paid subject-experimenters. Each was highly trained and served twice in each of 13 experimental conditions. The subject-experimenter alternation occurred within each 2-h experimental block and was counterbalanced across the 13 experimental days.

Apparatus

The apparatus consisted of a Scientific Prototype Model 323-G three-channel tachisto-scope, a Hunter electronic noise relay, and Hunter millisecond Klockcounter. One channel of the tachistoscope was used as a stimulus field and the other as a pre- and post-exposure field. The illumination intensities for the two fields were subjectively equated at a level high enough to permit easy viewing of the stimulus materials. The illumination level in the experimental room was slightly less than that of the experimental fields. A hand-held push-button permitted the subject to initiate each trial, and his vocal response activated a noise relay that terminated the trial.

Stimulus sets

The stimuli were squares varying in either size alone (S), brightness alone (B), size and brightness redundantly (SB-R), or size and brightness non-redundantly (SB-NR). All stimuli were centre mounted on 5×7 in ($12 \cdot 70 \times 17 \cdot 78$ cm) white tag board.

In order to provide a sufficient data range for a regression analysis, three set sizes (2, 4, 8) were employed within the S, B, and SB-R dimensional conditions and four (4, 6, 6′, 9) within the SB-NR condition. There are two ways of forming a non-redundant set of size six: 6 is the orthogonal combination of three size values and two brightness values; 6′ is the orthogonal combination of two size values and three brightness values.

The areas of the eight S stimuli (cut from Munsell paper N3/) varied by successive ratios of $1 \cdot 5$, the smallest being $0 \cdot 156$ cm^2 and the largest being $2 \cdot 669$ cm^2 ($S_1, S_2, \ldots, S_8$). The eight B stimuli were 1 cm^2 cut from Munsell paper varying in half steps from N8/ to N4$\cdot$5/ ($B_1, B_2, \ldots, B_8$). The SB-R set was formed by a one to one combination of size and brightness ($S_1B_1, S_2B_2, \ldots, S_8B_8$). The SB-NR set was the orthogonal combination of S_1, S_2, S_3, with B_6, B_7, B_8 ($S_1B_6, S_1B_7, S_1B_8, S_2B_6, \ldots, S_3B_8$). These values were selected so that the SB-R and SB-NR sets would have no stimuli in common thereby reducing transfer and response confusion effects.

For the S, B, and SB-R conditions, the two and four alternative sets were composed of the middle stimulus values. For the SB-NR condition, the four alternative set was the ortho-gonal combination of S_1, S_2, with B_6, B_7. The compositions of the two six-alternative sets have already been described.

The response names for the stimuli in S, B, and SB-R sets were the numbers one to eight preceded by the consonant "b": bun, boo, bee, bore, bive, bix, bev, bate. The names were chosen to provide a uniform stop consonant for the noise relay and to minimize the time required to learn the S-R associations. The names were assigned in order, e.g. S_1 = bun, S_2 = boo.

The response names for the non-redundant set were nine morphemes beginning with "b": bay, bet, biz, buk, ban, beg, bid, buz, bal. The method of set formation and the experimental design argued against using the S, B, SB-R names for this condition since the same subjects viewed all conditions. If the set had been numbered consecutively one to nine and then divided into a four or six alternative set, the response sets would have been a collection of non-consecuitve numbers, e.g. the four alternative response set would be one, two, four, five. Non-consecutive response sets have been shown to be less compatible and more difficult to learn than consecutively numbered sets (Fitts and Seeger, 1953). The use of non-numeric morphemes eliminated this difficulty. Both the numeric and non-numeric morpheme sets were tested for ease of learning with naive subjects prior to the experiment. The mean learning time and mean number of correct responses were almost identical.

Procedure

Each subject had four practice days prior to the experimental trials and received a total of 2 h of practice on the maximum alternative set for each condition. On the fourth day practice was arranged to avoid exposure to the condition that the subject would be tested on during the first experimental day. An experimental day consisted of two 2-h blocks separated by a minimum of 4 h. During each 2-h block, the subjects served 1 h as a subject and 1 h as the experimenter.

At the beginning of the first practice session, the subject was given instructions explaining the experimental situation. Then, the subject was shown the stimuli to be viewed in that session and told the responses to be associated with each. The subject was encouraged to practice the responses and given as much help as possible in learning them. Trials, with tachistoscopic presentation of the stimuli, were then given for 1 h.

The test session instructions provided a brief review of the experimental situation and designated the particular condition to be viewed. During each session, the subject viewed one size set from a particular dimensional condition for 144 trials with a brief rest every 18 trials. The first 36 trials functioned as a warm-up and were not included in the analyses.

The order of events on a given trial was as follows. The subject pushed the hand-held button which initiated the stimulus presentation and started the clock. The subject's vocal response turned off the stimulus and stopped the clock. The experimenter recorded the response and CRT on a prepared data sheet while reporting the correct response and CRT to the subject. Then, the experimenter reset the Klockcounter, changed the stimulus card in the tachistoscope, and said "set" as a cue to the subject that he could initiate the next trial whenever ready.

Results

The amount of information transmitted and median CRT calculated over the last 108 trials in each session for each subject are given in Table I. There were consistent differences between subjects and replications with subject 1 being faster than subject 2 and replication II performance better than replication I. However, since both subjects appeared to follow the same basic trend over both replications,
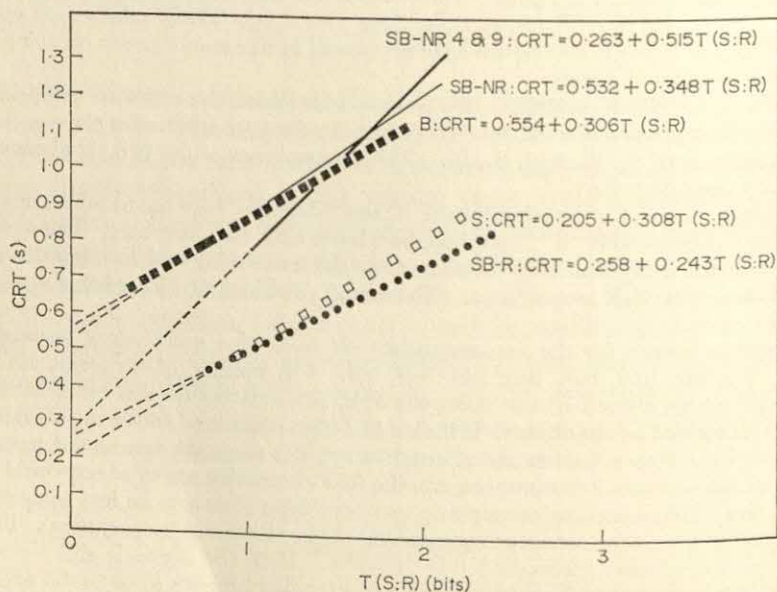

FIGURE 1. Linear regression lines and equations for CRT on T(S : R) (Experiment I).

the regression analysis of CRT on T(S:R) was done by pooling over subjects and replications for each dimensional condition. Since the residual variability for tests of the linear components contained the subject to subject and replication to replication variability, the tests can be considered conservative.

TABLE I

Median choice reaction time (CRT) in seconds and transmitted information (T(S:R)) in bits for each condition in Experiment I

Condition	Set size	Subject 1				Subject 2			
		Replication I		Replication II		Replication I		Replication II	
		T(S:R)	CRT	T(S:R)	CRT	T(S:R)	CRT	T(S:R)	CRT
Size	2	0·819	0·430	0·933	0·404	0·933	0·440	1·000	0·579
	4	1·388	0·569	1·749	0·559	1·415	0·794	1·395	0·854
	8	2·030	0·733	2·049	0·752	2·176	1·044	2·199	0·872
Brightness	2	0·444	0·615	0·689	0·636	0·697	0·758	0·601	0·754
	4	0·975	0·851	1·098	0·734	1·076	1·140	1·039	1·042
	8	1·669	1·078	1·853	0·934	1·821	1·140	1·537	1·097
Size-brightness redundant	2	1·000	0·428	1·000	0·466	0·933	0·556	0·933	0·508
	4	1·565	0·651	1·694	0·497	1·714	0·794	1·885	0·753
	8	2·359	0·851	2·387	0·638	2·247	0·926	2·322	0·899
Size-brightness non-redundant	4	0·989	0·751	1·192	0·658	1·239	1·137	1·374	0·972
	6	1·733	0·864	1·782	0·812	1·573	1·256	1·875	1·212
	6'	1·433	1·032	1·675	0·851	1·402	1·429	1·497	1·326
	9	2·049	1·152	1·973	0·955	1·902	1·545	2·036	1·513

The best fit least squares regression lines are shown for the SB-NR conditions: one, SB-NR, is fit to all SB-NR data points; the second, SB-NR 4 & 9, is fit to the 4 and 9 alternative data points excluding the 6 and 6' points. For the S, B, and SB-R conditions the linear coefficients, but no higher order components, were significantly different from zero (for S, $F = 16·6$, $df = 1, 10$, $P < 0·005$, for B, $F = 13·5$, $df = 1, 10$, $P < 0·005$, for SB-R, $F = 18·9$, $df = 1, 10$, $P < 0·005$. The linear component for the SB-NR condition was not significant ($F = 2·8$, $df = 1, 14$, $P > 0·10$). When a regression line was fit to the data points obtained with 4 and 9 SB-NR alternatives, a significant linear trend was found ($F = 5·8$, $df = 1, 6$, $P < 0·10$). The clustering of data points about values of transmitted information for 4 and 9 alternatives precluded testing higher order components.

The slopes for the S and B conditions were virtually identical (slope for S, 0·308 s/bit; slope for B, 0·306 s/bit). However, the intercepts differed greatly (intercept for S, 0·205 s; intercept for B, 0·554 s). The slope of the SB-R regression line was less than that of the S, B, and SB-NR regression lines. The SB-NR 4 & 9 regression lines slope was greater than that of the S, B, and SB-R conditions.

Experiment II

Although a slope (rate) difference was found between the single dimension and redundant bidimensional conditions in Experiment I, it was not substantial. The purpose of Experiment II was to determine if the reduced slope or increased rate for the SB-R condition was reliable. Only one single dimension condition, S, and

the SB-R condition were run since the results for the two single dimension conditions in Experiment I were so similar. A between-groups design was used instead of a within-subjects.

Method

Subjects

The subjects were 30 volunteers from an Introductory Psychology course at Indiana State University. All subjects had normal (or corrected) vision. Each subject served in three ½-h experimental sessions conducted on different days but within a 4-day period. Five subjects were assigned randomly to each of the six experimental conditions.

Apparatus

The apparatus was the same as in Experiment I except that a Scientific Prototype Model 800-E two-channel tachistoscope was used.

Stimulus sets

The stimuli were squares varying in either size (S) or size and brightness redundantly (SB-R). All stimuli were centre mounted on 4×5 in ($10 \cdot 16 \times 12 \cdot 70$ cm) white tag board. Three set sizes (2, 4, 8) were employed for both the S and SB-R conditions to provide a sufficient data range for a regression analysis.

Varying by successive ratios of $1 \cdot 32$, eight squares were cut from Munsell paper N6/; the smallest square being $1 \cdot 00$ cm^2 and the largest square being $7 \cdot 07$ cm^2. These stimuli were the set varying in size ($S_1, S_2, S_3, \ldots, S_8$). The SB-R set was formed by a one-to-one combination of size and brightness ($S_1B_1, S_2B_2, S_3B_3, \ldots, S_8B_8$). The smallest square was $1 \cdot 00$ cm^2 cut from Munsell paper N8/ and the largest was $7 \cdot 07$ cm^2 cut from Munsell paper N4·5/. The area was varied by successive ratios of $1 \cdot 32$ and the brightness was varied in Munsell half-steps from N8/ to N4·5/.

Set size 2 consisted of S_1, S_2 for the S condition and $S_1B_1, S_2B_2,$ for the SB-R condition, set size 4 consisted of S_1, S_2, S_3, S_4 for the S condition and $S_1B_1, S_2B_2, S_3B_3, S_4B_4,$ for the SB-R condition and similarly for set size 8.

Response names for the stimuli in the S and SB-R sets were the numbers one to eight preceded by the consonant "b": bun, boo, bee, bore, bive, bix, bev, bate. These names were assigned to the stimuli in ascending order, e.g. $S_1 =$ bun, $S_2 =$ boo.

Procedure

The procedure was nearly identical to that used in Experiment I. At the beginning of the first session, instructions concerning the nature of the experiment were given and the assignment of response names to stimuli was described. Then subjects practiced naming the stimuli as they were shown in the tachistoscope. A series of ·120 practice trials was then given with subjects instructed to respond as fast and as accurately as possible. Instructions were briefly reviewed at the beginning of the second and third sessions. A trial began when the subject pushed the hand-held button which brought the stimulus into view and started the clock. The subject's vocal response removed the stimulus from view and stopped the clock. The experimenter then recorded the response and CRT on a data sheet and gave the subject error and CRT feedback. Then, the experimenter reset the clock to zero, put a new stimulus card into the tachistoscope and said "set" as a cue to the subject that he could initiate the next trial when ready. The second and third sessions began with 24 practice trials followed by 96 test trials. In all sessions rest periods were given after blocks of 32 trials.

Results

The amount of information transmitted and median CRT calculated over the last 96 trials of the third session for each subject are shown in Figure 2 as well as the best fit linear regression lines for the S and SB-R conditions. Data from two
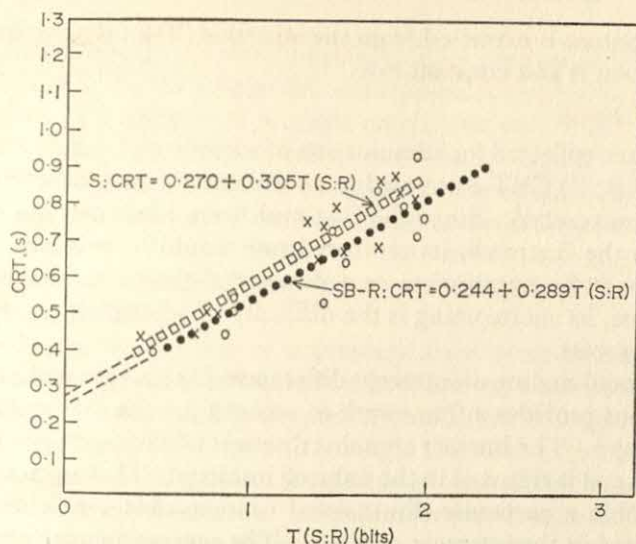
FIGURE 2.   Data points (Xs for S, os for SB–R), linear regression lines and equations for CRT on T(S : R) (Experiment II).

subjects, one in the S 2 alternative and one in the SB-R 2 alternative condition, were discarded since their scores deviated greatly from all other subjects in the condition.

For both the S and SB-R conditions the linear components and no higher order components of the regression of CRT on T(S : R) were significant (for S, $F = 72 \cdot 0$, $df = 1, 12$, $P < 0 \cdot 0005$, for SB-R, $F = 51 \cdot 6$, $df = 1, 12$, $P < 0 \cdot 0005$. The slopes of the SB-R and S conditions were nearly identical (slope for S, $0 \cdot 305$ s/bit; slope for SB-R, $0 \cdot 289$ s/bit).

Discussion

The finding of significant linear relationships between CRT and T(S : R) for the S, B, and SB-R conditions clearly extends the empirical support for Hick's law, i.e. that the rate of gain of information is a constant. Moreover, the groundwork is laid for separating various components in stimulus identification processes as a function of stimulus dimensionality.

The discovery of nearly identical slopes of the regression lines for the unidimensional and redundant bidimensional conditions is noteworthy. If one value for the slope had been found for the unidimensional condition and another for the redundant condition, then we would have argued that a constant capacity to identify unidimensional stimuli exists for the rate of information processing measured in bits/s as well as in bits/stimulus (Miller, 1956). However, the finding of a constant slope suggests that different stages of the identification process are being measured. One stage has a limited capacity in bits/stimulus or amount of information transmitted, in the sense of channel capacity to which Miller referred, the output at this stage is a function of dimensionality. In the present experiments, T(S : R) increased for the redundant conditions which is consistent with previous findings. The slope measure then taps another stage of processing which is independent of dimensionality and where capacity is limited in bits/s, i.e. once the

informational content is extracted from the stimulus $(T(S:R))$, further processing of this information is at a constant rate.

Unidimensional stimuli

If data had been collected for stimulus sets of a single size, e.g. eight alternatives, and the ratio $T(S:R)/CRT$ computed as an estimate of rate, different conclusions would have been reached. Since the S stimuli were identified more quickly and accurately than the B stimuli, its estimated rate would be greater. Although the ratio measure may be appropriate as a direct real-time measure of information transmission rate, its shortcoming is the difficulty of separating components of the identification process.

Doherty's (1968) finding of intercept differences for his easy and hard discriminability conditions provides a framework to account for the difference between the S and B intercepts. The intraset stimulus discriminability is more pronounced for the size stimuli and is reflected in the reduced intercept. That is, stimulus variation or selection within a particular dimensional context adds a constant to the CRT which is reflected in the intercept constant. The regression analysis permits us to separate the information processing *per se* from the effects of other stimulus characteristics which may be difficult to equate across dimensional conditions.

Redundant stimuli

The results of Experiments I and II taken together indicate that there is no difference in the rate of processing the unidimensional and redundant bidimensional stimuli. Since the CRTs for the SB-R condition were of the same order as those for the S condition in Experiments I and II, it appears that the subjects responded primarily to stimulus properties associated with size. It would seem as if the subjects' strategy was to identify the size component of the redundant stimulus and to ignore the redundant brightness component. However, the S and SB-R regression lines were not congruent as one would expect if the size dimension were completely dominant: there were fewer errors and consequently an increase in transmitted information as opposed to faster CRTs for the SB-R condition in Experiments I and II.

The increase in $T(S:R)$ could be a function of (1) intraset discriminability, (2) parallel processing, or (3) dimensional non-independence. If the addition of redundant dimensions functions solely to increase intraset discriminability, we would predict, based on the argument presented above for the difference between the S and B intercepts, a change in intercept, rather than slope. The data from Experiment II lend some support for this explanation; however, without both unidimensional conditions in the same experiment, it is impossible to reach any valid conclusions. A second way to obtain the reduction in errors associated with the increase in transmitted information is for the subjects to process simultaneously the stimulus values along both dimensions and then make a decision according to some algorithm (see Eriksen and Hake, 1955, for an example of a combinatorial algorithm for predicting responses to redundant bidimensional stimuli from the responses to the unidimensional components). Since the CRTs for the SB-R condition were the same or slightly faster than for the unidimensional condition, this explanation is tenable assuming that the processing of information on

component dimensions occurs simultaneously rather than sequentially. The third explanation focuses on the particular dimensional combination. When varied simultaneously in the context of a single patch, size and brightness may not be perceptually independent. Intuitively, one would expect "perceptual non-independence" to lead to a decrease in transmitted information. However, Kaufman and Levy (1971), in an analysis of various effects of combining dimensions, showed that the consequence of most forms of "perceptual non-independence" is to *increase* transmitted information. Any one or a combination of these interpretations could account for the data.

The generality of the finding of a constant slope must be extended to other redundancy conditions. A linear correlated redundant combination of values on component dimension was used in the present experiment. Lockhead (1970) found greater increases in $T(S:R)$ from the single dimension conditions for those redundant combinations which he described as "sawtooth" than for the linear correlated redundant combinations. Additional research is needed with different types of redundant sets to determine whether or not the increased transmitted information reported by Lockhead (1970) is accompanied by changes in intercept and/or slope.

Non-redundant stimuli

Comparison of the SB-NR 4 & 9 results showed the inclusion of the 6 and 6' sets had the effect of decreasing the slope of the line. The shift in slope is possibly artifactual. The relatively smaller slope for the overall SB-NR conditions could reflect either a true difference in processing rate or a noise component due to a different relationship evidenced by the failure to find a significant linear component.

The results for the SB-NR condition should be interpreted cautiously. Although the reasoning behind the choice of different response schemes seemed sound, subjects employed one response set 75% of the time and the other set 25% of the time. This lack of practice on the SB-NR response set may have made it more difficult to employ. In addition, the SB-NR set lacked the natural ordering present in the other sets.

Based on the differences in response sets, a tentative explanation of the increased slope of the SB-NR condition is possible. (The regression line SB-NR 4 & 9 is used as the basis of argument since it had a significant linear component.) For the S, B, and SB-R conditions, in which there were the same natural orderings of stimuli and responses, the slopes were the same. In the SB-NR condition different responses were used and the natural ordering did not maintain reducing the stimulus-response compatibility. Assuming that the same stages of processing apply to the SB-NR condition as hypothesized above for the S, B, and SB-R conditions, dimensionality effects *per se* on discriminability and on the number of stimuli that can be identified would affect CRT and $T(S:R)$ independent of the slope. The locus of the SB-NR 4 & 9 intercept, between the S and B intercepts, is consistent with what we would have expected if discriminability and dimensionality affect the intercept and not the slope. The response code difference would then account for the increased slope, suggesting furthermore, that the stimulus-response code is an important determinant of rate at this processing stage.

The average CRT was faster and $T(S:R)$ larger for the stimuli in the 6 set, which consisted of three values of size paired with two values of brightness, than in the 6' set. The subject was responding to interdimensional differences as well as the intradimensional load. As in the SB-R conditions in Experiment I, the size dimension was relatively dominant. To speculate, the analysis of orthogonal combinations of stimulus values in which the numbers of values on component dimensions are unequal may provide some insights into the underlying processes. It may be easier to separate effects of information load and dimensional interactions when the number of values on the component dimensions is different.

Although the results in the present study for the non-redundant condition are not conclusive, Garner and Morton (1969) and Kaufman and Levy (1971) have pointed out that this condition is fundamental to the analysis of the combination of perceptual dimensions. Additional research is needed not only to compare the efficiency and manner of processing of redundant and non-redundant bidimensional stimuli but also to do so within the context of an analysis of different dimensional combinations. For example, the effects of stimulus "integrality" (Lockhead, 1966; Garner and Felfoldy, 1970) on the stages of information processing should be studied. It might be possible to relate the concept of "integrality" to the components of processing Kaufman and Levy (1971) present in their general analysis.

Conclusions

Analysis of the intercepts and slopes of the regression of CRT on $T(S:R)$ has provided a vehicle for separating stages of the identification process. The method of analysis is in the same vein as the one Sternberg (1969) proposed for studying memory scanning using reaction time measurements. Including an error sensitive dependent measure, $T(S:R)$, may have the advantage of permitting the investigation of processing stages which do not affect CRT directly.

The effects of stimulus dimensionality on the rate of gain of information as measured from the slope of the regression of choice reaction time on transmitted information may depend primarily on stimulus-response relationships. Given comparable stimulus-response pairings, the "identification" *per se*, i.e., attaching the particular identifying response to the stimulus which has passed through a stage with a bits/stimulus capacity as Miller (1956) noted, is executed at a constant rate in bits/s as Hick (1952) noted. Additional research is necessary to test the generality of the present findings and to determine more fully the variables affecting the different stages.

References

ATTNEAVE, F. (1959). *Applications of Information Theory to Psychology.* New York: Holt.

BEEBE-CENTER, J. G., ROGERS, M. S. and O'CONNELL, D. N. (1955). Transmission of information about sucrose and saline solutions through the sense of taste. *Journal of Psychology,* **39**, 157–60.

DOHERTY, M. E. (1968). Information and discriminability as determinants of absolute judgment choice reaction time. *Perception and Psychophysics,* **3**, 1–4.

ERIKSEN, C. W. and HAKE, H. W. (1955). Multidimensional stimulus differences and accuracy of discrimination. *Journal of Experimental Psychology,* **50**, 153–60.

FITTS, P. M. and SEEGER, C. M. (1953). S-R compatability: Spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology,* **46**, 199–210.

GARNER, W. R. and FELFOLDY, G. L. (1970). Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology,* **1**, 225–41.

GARNER, W. R. and MORTON, J. (1969). Perceptual independence: definitions, models and experimental paradigms. *Psychological Bulletin,* **72**, 233–59.

HICK, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology,* **4**, 11–26.

KAUFMAN, H. and LEVY, R. M. (1971). An analysis of the combination of perceptual dimensions. *Psychological Bulletin,* **75**, 369–78.

LEVY, R. M. (1965). The effects of stimulus dimensionality on information processing. *The Proceedings of the 73rd Annual Convention of the American Psychological Association,* Washington, D.C.: APA, 47–48.

LOCKHEAD, G. R. (1966). Effects of dimensional redundancy on visual discrimination. *Journal of Experimental Psychology,* **72**, 95–104.

LOCKHEAD, G. R. (1970). Identification and the form of multidimensional discrimination space. *Journal of Experimental Psychology,* **85**, 1–10.

MILLER, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review,* **53**, 81–97.

NYE, P. W. (1962). Aural recognition time for multidimensional signals. *Nature,* **196**, 1282–3.

POLLACK, I. (1953). The information of elementary auditory displays II. *Journal of the Acoustical Society of America,* **25**, 765–9.

POLLACK, I. and FICKS, L. (1954). Information of elementary multidimensional auditory displays. *Journal of the Acoustical Society of America,* **26**, 155–8.

STERNBERG, S. (1969). Memory scanning: Mental processes revealed by reaction-time experiments. *American Scientist,* **57**, 421–57.

WHAT IS BOTH NECESSARY AND SUFFICIENT TO MAINTAIN AVOIDANCE RESPONDING IN THE SHUTTLE BOX?

RICHARD KATZEV†

Reed College, Portland, Oregon

To learn what maintains the frequency of shuttle box avoidance responses, male rats from the Berkeley S1 strain, after 200 trials of standard discriminative avoidance training, were given 100 additional trials under one of four different conditions. Responding at the maximum rate was maintained when animals performed under the training conditions or when responses continued to terminate the warning signal immediately, even though shock was never given for failing to respond. In contrast, avoidance responding was reduced markedly if, and only if, trials were given in which the signal ceased to terminate immediately (i.e. it shut off either well before or well after a response). This decrement occurred even though avoidance responses continued to avert shock. Thus, under the conditions of this experiment prompt signal offset was *both* necessary and sufficient to maintain the occurrence of well-established shuttle box avoidance responses.

Introduction

What maintains the frequency of discriminative avoidance responses? In two recent reports (Bolles, Stokes and Younger, 1966; Bolles and Grossen, 1969) it has been claimed that termination of the warning signal (CS) is far less important in controlling such behavior than it is usually assumed to be. The evidence for this claim is drawn from a set of studies in which the relationship between avoidance responding and CS termination was varied in several ways. In one set of studies Bolles and Grossen (1969) report that animals readily learn to perform avoidance responses, even though each such response is never followed by immediate termination of the CS. In another set, Bolles *et al.* (1966) report that even when responses do terminate the CS immediately, animals fail to learn as long as such responses do not avoid shocks. As a result, these authors conclude that prompt CS termination is neither necessary nor sufficient to strengthen and maintain avoidance responses.

However, before we can accept this proposition, it is important to point out that each of these studies, as well as other recent investigations of this contingency (D'Amato, Fazzaro and Etkin, 1968; Cassady, Cole, Hall and Williams, 1971) have studied the *initial* acquisition of avoidance behavior. It is essential, though, to distinguish this behavioral process and the mechanisms governing it from the one that occurs once the response has been stabilized at a terminal rate. There is no *prima facie* reason to assume that the variables which control initial learning are the same as those controlling the *maintenance* of the response after it has so stabilized.

In fact, there is evidence from a study by Verhave (1959), as well as a number of other studies on the extinction of avoidance responses (e.g. Solomon, Kamin and Wynne, 1953; Robinson, 1961; Katzev, 1967), which indicate that the CS termination contingency does play a critical role in maintaining the frequency of well-learned avoidance responses.

Accordingly, the following experiment was designed to test the generality of the Bolles *et al.* claim for the maintenance of avoidance responses, once they have been established. After 200 trials of standard avoidance training, the animals were given 100 additional trials in which the contingency for presenting either shock or CS termination was altered. The shock contingency was changed in one condition by no longer presenting it when the animals failed to respond. The CS contingency was changed by ceasing to terminate it immediately after each avoidance response. In one condition it terminated well *before* the response; in the other condition its termination was delayed until well *after* the response. In both of these latter conditions avoidance responses continued to avert shock. The behavior of animals under each of these conditions was evaluated against a baseline occurring when animals were continued under the original training conditions. From the outcome of such comparisons it is possible to test the hypothesis that prompt signal offset is both necessary and sufficient to maintain well-established avoidance responses.

Method

Subjects

Thirty experimentally naive male rats from the Berkeley S1 strain (descendents from the Tryon Maze Bright stock) were employed in the study. They were between 127 and 154 days of age at the start of training and were housed in separate cages, where they had free access to food and water.

Apparatus

The experimental chamber was a shuttle box without a barrier or door modelled after the one used by Kamin (1956). Its interior dimensions were: length, 91·4 cm; width, 13·9 cm; height, 12·7 cm. The CS consisted of a compound buzzer-light stimulus. Light entered the box from bulbs mounted 39·4 cm from the two end sides. The buzzer was mounted outside the box midway between the two compartments. Shock (US) with a current flow of 2 mA was delivered through a silent electronic scrambling unit (Markowitz and Saslow, 1964) to the grid floor bars of each side of the box separately.

Procedure

Each animal was tested in a block of 25 trials for 12 consecutive days. The intertrial interval during each such session ranged from 40 to 80 s with a mean of 1 min.

Acquisition. Each animal was given a 3 min adaptation period prior to the start of the first trial. The CS–US interval for trials during this phase of the experiment was 5 s. Whenever the rat crossed into the opposite compartment within 5 s after the start of a trial, the CS terminated immediately and shock was not given. If the rat failed to cross within 5 s, shock was delivered to the side of the box he was in and both the CS and US remained on until the animal made a shuttle response, whereupon both were promptly terminated.

Maintenance. Following these trials, four groups, matched on the basis of their performance during the final 100 training trials, were given 100 additional trials. The following conditions were employed. (1) Training Control (N = 6) continued for an additional 100 trials under conditions identical to those which held during the first 100 trials. (2) Prompt Signal Offset (N = 8) was treated similarly to the Training Control, except that shock was

no longer delivered on trials in which an animal failed to respond. If the animal did not respond within 60 s, the CS was arbitrarily terminated by the experimenter. However, whenever the animal did respond, the CS terminated immediately. (3) Delayed Maintenance ($N = 8$) continued to receive shock if, and only if, a CR did not occur within 5 s after the start of a trial. The termination of the CS was, however, delayed for 20 s after the first response which occurred during the trial, regardless of its latency. During this 20-s period the rat was free to cross back and forth between compartments; such behavior did not affect the end of the trial, which always came 20 s after the first crossing. (4) Trace Maintenance ($N = 8$) also continued to receive shock if, and only if, a response did not occur within 5 s. In this condition, the CS was presented according to a trace procedure whereby it terminated before the animal could respond. To accomplish this the duration of the CS was 0·5 s on all trials during this phase of the experiment.

Results

Within the first 100 training trials the animals gradually approached the maximum (i.e. 22 or more CRs in a block of 25 trials) rate of avoidance responding and continued to perform at that rate throughout the remaining 100 training trials. Figure 1 depicts the level of avoidance responding under each condition, when the animals were switched to the maintenance phase of the study. There was no decline in the level of avoidance responding when the animals continued under the Training Control conditions. Figure 1 also shows that animals switched to conditions of Prompt Signal Offset continued to respond at a rate which was virtually indistinguishable from that of the Training Control. Although these two groups did not differ significantly in the overall frequency of CRs throughout the maintenance sessions, they did differ in terms of the latencies of these responses. Figure 2 plots the mean latency of CRs under each of these conditions during the last 100 trials of acquisition and the subsequent 100 maintenance trials. There is considerable overlap in the distribution of latencies during the final 100 training trials and the groups do not differ on this measure. However, as Figure 2 indicates, by the last block of maintenance trials, animals were responding much more quickly under conditions of Prompt Signal Offset than they were under the Training Control conditions. Separate Mann Whitney U-tests indicated that the mean CR latency on this last block of trials, as well as the overall mean CR latency throughout the maintenance sessions, was significantly less under Prompt Signal Offset than it was under Training Control ($U = 7$, $P < 0.02$; $U = 9$, $P < 0.03$, respectively).

Figure 1 shows that switching to conditions of Delayed Maintenance produced an immediate and sustained reduction in the rate of avoidance responding to a level well below that occurring under Training Control and Prompt Signal Offset. Under Delayed Maintenance the mean percentage of avoidance responses during these trials was 28·3 (range 11–52). This contrasts with a mean of 92·2 (range 78–99) in the Training Control and a mean of 91·6 (range 74–99) under Prompt Signal Offset. The fact that there is no overlap in the distribution of CRs between each of these latter two groups and Delayed Maintenance highlights the striking and uniform decrement in performance observed under this condition. Although six of the animals in Delayed Maintenance had stabilized at a moderate rate of avoidance responding, two others were responding at a minimum rate by the end of
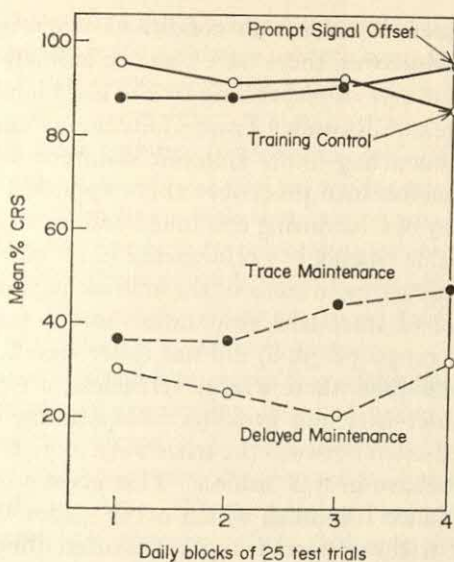
FIGURE 1. Mean percentage of avoidance responses under each experimental condition during each block of maintenance test trials.

these sessions. This meant that they were getting shocks for failing to respond on at least 20 out of the 25 daily trials.

As Figure 1 indicates there was also an immediate reduction in the rate of avoidance responding when the animals were switched to Trace Maintenance conditions. The overall percentage of avoidance responding in this group was 42·3 (range 20–70). Again there was no overlap in the distribution of CRs between this
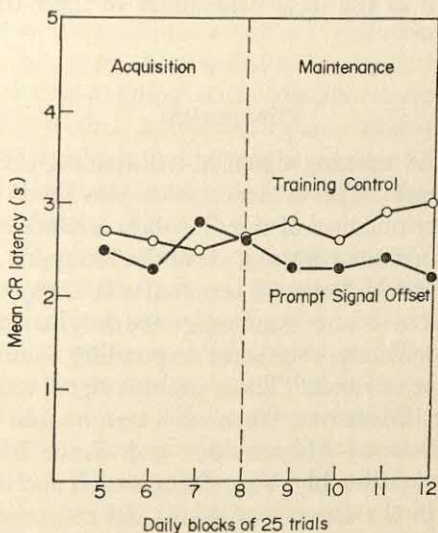
FIGURE 2. Mean latency of avoidance responses in each block of the last 100 acquisition trials and the 100 maintenance trials for Prompt Signal Offset and Training Control.

condition and that observed under both conditions where responding continued at the maximum rate. Moreover, there were also two animals in this condition who dropped to the minimum rate of responding by the last block of trials. Lastly, the overall mean percentage of CRs under Trace Maintenance conditions did not differ significantly from that occurring under Delayed Maintenance ($U = 17$, $P > 0.05$).

Under the Trace Maintenance procedure there appeared to be a slight overall increase in the frequency of CRs during continued maintenance sessions. However, the meaning attached this change in performance is obscured by the high rate of inter-trial responding occurring in some of the animals in this condition. Although the average daily level of inter-trial responding under the Trace Maintenance condition (mean 19.4; range 3.8–39.8) did not differ significantly from any of the other experimental conditions, there was, nevertheless, a high positive correlation ($\rho = 0.91$) between inter-trial and avoidance responding in this group. Thus animals who responded often between the trials were more likely to make CRs than animals who did not behave in this fashion. This poses a difficult problem in the interpretation of avoidance responses which occur under this procedure. When the frequency of inter-trial responses is high, it is often difficult to tell if a response occurring after the start of a trial is under the control of the CS, or if it is an inter-trial response which occurs quite by chance within the CS-US interval. This problem exists not only under the trace procedure, where it has also been noted by other investigators (Kamin, 1954; Black, 1963) but also in other situations, such as the running wheel (see Bolles *et al.*, 1966, Experiment 4) where the level of spontaneous responding is high. Because it is often difficult to distinguish inter-trial and avoidance responses in studies employing these situations, it is desirable to view their outcomes with considerable caution. In order to be confident that the reported CR rate is a true index of discriminative responding it is essential to be able to demonstrate that this rate is independent of the overall probability of inter-trial responding, as well as the local probability of inter-trial responding at the time the trial begins.

Discussion

What is the role of the warning signal in maintaining the occurrence of shuttle box avoidance responses? The evidence from this study flatly contradicts the claim that immediate termination of the CS plays a relatively minor role in reinforcing such behavior. On the contrary, all of the foregoing evidence supports the belief that the manner in which the CS terminates is a major factor controlling the frequency of shuttle box avoidance responses once they become established. As long as the CS shut off immediately, avoidance responding continued at the maximum rate, even in the absence of shock. Thus, prompt signal termination was *sufficient* to maintain responding. Moreover, when such termination was omitted, as it was under conditions of Delayed Maintenance and Trace Maintenance, avoidance responding weakened considerably, even though each such response continued to avoid shock. Thus, in the absence of additional response contingent feedback stimuli (e.g. Bolles and Grossen, 1969) prompt signal termination was *necessary* to maintain responding. In short, all of the evidence in this study supports the

proposition: *In rats of the Berkeley S1 strain, prompt signal termination is both necessary and sufficient to maintain the maximum rate of shuttle box avoidance responding.*

Furthermore, there is little support in this study for the belief that shock avoidance is the principal variable controlling the maintenance of such behavior. Each time the animal made a quick response, under both conditions of Delayed Maintenance and Trace Maintenance, shock was avoided on that trial. In spite of such trials and in spite of frequent experience with that contingency during earlier training trials, these animals did not continue to respond at a rate even remotely close to their former level. Because these rats fell significantly below the maximum rate, it apears that: *In rats of the Berkeley S1 strain, shock avoidance is not sufficient to maintain the frequency of shuttle box avoidance responses.*

However, this conclusion should be carefully qualified by describing somewhat more exactly the actual rate of responding observed under these conditions. It is important to point out, as Sidman (1966) has done for the unsignaled avoidance situation, that responding under the discriminative avoidance procedure often stabilizes at several different rates. Thus, in this situation, it is possible to distinguish, albeit in a somewhat arbitrary fashion, between the maximum rate (i.e. 22 or more CRs in a block of 25 trials) and the minimum rate (i.e. 5 or fewer CRs in block of 25 trials), as well as the moderate rate which differs significantly from either one or both of these extremes. Although shocking an animal if, and only if, he fails to respond is not sufficient to maintain responding at the maximum rate in Berkeley S1 rats, it would be erroneous to conclude thereby that this contingency does not control behavior. On the contrary, it is clear from the evidence in this study that such a shock schedule is sufficient to maintain responding at a moderate rate.

Interpretations of the experimental facts of avoidance behavior have often been clouded by failing to distinguish between these rates. For example, Bolles *et al.* (1966) conclude from their investigations that "there is reason to doubt whether CS termination serves as the principal source of reinforcement for the CR". Yet, in none of their experiments on the acquisition of shuttle box avoidance responses did these authors report responding at the maximum rate *in the absence* of that contingency. Inspection of their tables and figures indicates that in each instance where the animals performed without benefit of prompt CS offset, responding was significantly below the level occurring under comparable conditions where the CS did terminate immediately. Thus, it would have been more exact to conclude that although prompt CS offset was not necessary to establish the moderate rate of shuttle box avoidance responding, it was necessary to establish responding at the maximum rate. Moreover, it is also clear from this same report that prompt CS termination was not sufficient to establish avoidance behavior. Again, however, one should qualify this conclusion by noting that it holds only for the maximum rate. With the exception of a condition where *both* the avoidance and escape contingencies were removed from the training procedure, avoidance responding always occurred at least at the moderate rate whenever prompt signal offset was present. Again, it would have been more exact to conclude that although prompt signal offset was not sufficient to establish the maximum rate of avoidance responding, it was, as Kamin (1956) had noted earlier, sufficient under several conditions

to produce shuttle box avoidance responding at a moderate rate. In general then, it is somewhat misleading to speak of the way this or any other contingency strengthens, maintains or acts as the "principal source of reinforcement" for avoidance learning, unless, at the same time, these rates of responding are explicitly distinguished. Although a contingency may not control responding at one rate, it may do so at another.

It is clear from the evidence in this study that the CS termination contingency does play a decisive role in maintaining shuttle box avoidance responses at the maximum rate. This finding may be interpreted in the light of the more general "discrimination hypothesis" (e.g. D'Amato et al., 1968). Thus, as pointed out elsewhere (Katzev, 1967) most of the trials which occur under Prompt Signal Offset are indistinguishable from the majority of the previous acquisition trials. As a result, it is virtually impossible for the animal to discriminate between these two kinds of trials and, therefore, he is quite likely to continue to respond for a rather indefinite period as long as such stimulation is present. Furthermore, omitting trials of prompt CS offset under Trace Maintenance and Delayed Maintenance conditions involves a marked change in the stimulus conditions which held during acquisition. The *immediate* decrement in performance which this change produced in both these groups suggests that omitting prompt CS offset not only removes an important source of reinforcement, but, at the same time, suddenly changes the discriminative function of the CS which may, in itself, weaken behavior.

These outcomes extend to the maintenance of avoidance responses the findings of several investigators (e.g. Kamin, 1956; Black, 1963) that prompt signal offset is necessary to establish shuttle box avoidance responses at the maximum rate. They also confirm Verhave's (1959) finding in the only other reported study which manipulated the CS termination contingency following extended acquisition sessions. Verhave reports that when animals were switched to conditions of delayed maintenance, performance on a wheel during avoidance response became extremely erratic, with the majority of animals failing to return to their steady state responding at the maximum rate. Lastly, the findings reported herein agree with the widely reported (e.g. Solomon et al., 1953) observation that following the acquisition of an avoidance response, trials of prompt signal offset continue to strengthen behavior and appear under certain conditions capable of maintaining the occurrence of short latency responses indefinitely. At the same time, they are consistent with the findings of several recent reports (Soltysik, 1960; Owen, 1963; Katzev, 1967) which indicate that avoidance responses can be rapidly and reliably weakened to the minimum rate once that contingency is removed.

In short, once an avoidance response becomes established there are a variety of conditions under which prompt signal termination is both necessary and sufficient to maintain it. This conclusion provides little support for those who have claimed that the CS is no more effective than other informational stimuli in controlling the occurrence of avoidance behaviors. On the contrary, on the basis of the evidence reported and reviewed herein, there is little doubt that the CS exerts considerably more influence over the maintenance of such behaviors than is implied by this formulation.

References

BLACK, A. H. (1963). The effects of CS–UCS interval on avoidance conditioning in the rat. *Canadian Journal of Psychology*, **17**, 174–81.

BOLLES, R. C. and GROSSEN, N. E. (1969). Effects of an informational stimulus on the acquisition of avoidance behavior in rats. *Journal of Comparative and Physiological Psychology*, **68**, 90–9.

BOLLES, R. C., STOKES, L. W. and YOUNGER, M. S. (1966). Does CS termination reinforce avoidance behavior? *Journal of Comparative and Physiological Psychology*, **62**, 201–7.

CASSADY, M., COLE, M., HALL, M. and WILLIAMS, T. (1971). The role of the CS and its relationship to the CR in avoidance conditioning. *Learning and Motivation*, **2**, 12–25.

D'AMATO, M. F., FAZZARO, J. and ETKIN, M. (1968). Anticipatory responding and avoidance discrimination as a factor in avoidance conditioning. *Journal of Experimental Psychology*, **77**, 41–7.

KAMIN, L. J. (1954). Traumatic avoidance learning: the effects of CS–US interval with a trace conditioning pocedure. *Journal of Comparative and Physiological Psychology*, **47**, 65–72.

KAMIN, L. J. (1956). Effects of termination of the CS and avoidance of the US on avoidance learning. *Journal of Comparative and Physiological Psychology*, **49**, 420–4.

KATZEV, R. (1967). Extinguishing avoidance responses as a function of delayed warning signal termination. *Journal of Experimental Psychology*, **75**, 339–44.

MARKOWITZ, H. and SASLOW, M. A. (1964). A reliable silent electronic shock scrambler. *Journal of the Experimental Analysis of Behavior*, **7**, 267–8.

OWEN, S. (1963). The effect on avoidance response extinction in rats of CS continuation and emotional constitution. *Journal of Genetic Psychology*, **103**, 147–51.

ROBINSON, H. (1961). Persistence of a response in the apparent absence of motivation. *Journal of Experimental Psychology*, **61**, 480–8.

SIDMAN, M. (1966). Avoidance behavior. In HONIG, W. K. (Ed.) *Operant Behavior: Areas of Research and Application*. New York: Appleton-Century-Crofts.

SOLOMON, R. L., KAMIN, L. J. and WYNNE, L. C. (1953). Traumatic avoidance learning: The outcomes of several extinction procedures with dogs. *Journal of Abnormal and Social Psychology*, **48**, 291–302.

SOLTYSIK, S. (1960). Studies on the avoidance conditioning: II. Differentiation and extinction of avoidance reflexes. *Acta Biologiae Experimentalis*, **20**, 171–82.

VERHAVE, T. (1959). Permanence of effect produced by delayed termination of warning stimulus in an avoidance situation. *Psychological Reports*, **5**, 31–8.

VISUAL PERSISTENCE AS MEASURED BY REACTION TIME

GARY G. BRIGGS

Department of Psychology, Duke University

MARCEL KINSBOURNE

Division of Pediatric Neurology, Duke University Medical Center

Latency of reaction to onset of a visual display was subtracted from latency of reaction to offset. Persistence was defined as difference between the two latency values. Persistence was inversely related to stimulus duration and was comparable for monoptic presentation and for presentation of the first half of the stimulus duration to one eye and the second half to the other. A power function described the relation between persistence and stimulus duration. The possible effects of central intermittency on this type of reaction time measure are discussed.

Introduction

Briefly presented visual information remains available to the subject for a period greater than the duration of the physical stimulus (Sperling, 1960). A duration of short term visual storage (STVS) of approximately 250 ms has been inferred from measurement of the number of items which subjects are able to process following the offset of a brief visual stimulus (Sperling, 1960). Averbach and Coriell (1961) and Keele and Chase (1967) arrived at similar estimates by varying the time between stimulus offset and presentation of a marker specifying recall of a particular single item. In none of these studies was stimulus duration reported to affect the duration of STVS. Other investigators have explicitly equated STVS with phenomenal persistence of the visual stimulus (Neisser, 1967).

A cross modal temporal matching paradigm in which the subject's task is to synchronize an auditory event with onset or offset of the visual stimulus has been employed by Haber and Standing (1969b) and by Efron (1970). Efron reported that for stimulus durations less than 130 ms persistence was inversely related to stimulus duration. No persistence was reported for stimulus durations longer than 130 ms. At all durations, persistence was much less than the 200 ms reported by Haber and Standing who also found persistence with stimulus durations as long as 1 s. This method measures persistence under conditions of cross modal shift in attention from the visual display to the auditory signal at or following offset. This attentional shift could introduce uncertainty into the subject's attempt at synchronization, or even suppress persistence, which might otherwise have endured for a longer period of time.

In one of several methods used by Haber and Standing (1969a) a white mask with a slit in it was briefly oscillated in front of a simple black on white test stimulus. The on to off ratio for the test stimulus was approximately 1 : 20. Average on time

was approximately 15 ms per cycle. Light-adapted subjects were required to report whether the stimulus was continuous or discontinuous over time. No information extraction was required. Persistence was approximately 300 ms. Similar results at the same exposure duration and on-off ratio were obtained when the pattern was presented tachistoscopically. The test field alternated with an adapting field of equal luminance. Decreasing luminance of both the fields resulted in an increase in persistence. Changing to dark pre- and post-exposure fields increased persistence to approximately 400 ms. In their reports of continuity the subjects may have been over-estimating or over-reporting the continuity of the stimulus in response to the implicit demand characteristics of the experimental situation.

A measure of the persistence of brief visual stimuli can be obtained by using Donders' method. The mean of a series of RTs to the onset of the stimulus can be subtracted from the mean of a series of RTs to the offset of a stimulus, resulting in a measure of mean persistence of the visual stimulus for each subject. To ensure that information is extracted and encoded during the stimulus presentation, subjects were given a supraspan display.

This design was applied to four questions regarding visual persistence: (1) Is apparent duration different when the stimulus provides little information (a binary decision—present or absent) as compared to when much information (letters) must be extracted while subjects are also attending to the temporal characteristics of the stimulus? (2) Are subjects unaware of the transition from physical stimulus to STVS as Neisser (1967) suggests? (3) What is the effect on persistence of monoptic as compared to dichoptic stimulation? A peripheral locus of persistence should give rise to a persistence which is a function of the duration of stimulation to a single eye. A central locus should give rise to a persistence the duration of which is a function of the total duration of stimulation to both eyes. (4) What is the effect of exposure duration on persistence? If persistence has adaptive significance, it might prolong a brief visual event until as much information as possible has been extracted. Thus, persistence may be greater at shorter durations of the stimulus where the limitation on output is still due to limited encoding rather than to the capacity of the short-term memory store.

Method

Subjects

The subjects were two males and one female ranging in age from 23 to 27 years. All had normal vision. All had extensive prior experience in reaction time studies. Subjects were paid for their services.

Apparatus

A modified four-field stereo tachistoscope (Polymetric Model V-1263) was used to present the stimuli. The lamp drivers were operated by square waves from two modified Tektronix Type 161 pulse generators in which larger capacitors had been installed to provide the necessary pulse widths. The lamps (Sylvania F4T5CW) were placed in a constant field of 540 VAC to maintain ionization and were continuously heated with 6·3 VAC. Firing was accomplished by 170 VDC pulses switched by mercury-wetted contact relays (Potter and

Brumfield JML-1110-81). Luminance as determined with a MacBeth illuminometer was approximately 4 ft.L.

The pulses which operated the lamp drivers also closed and locked an identical relay which operated a solid state digital stop clock (Electronics Research Co. 2623) if and only if a telegraph key were closed. Releasing the key stopped the clock. The same clock was used to calibrate the pulse duration to a level of accuracy of ± 1 ms. Synchronization of multiple pulse generator applications for dichoptic presentations was accomplished with a Tektronix dual beam storage oscilloscope (Type 146). Estimated alignment error was ± 2 ms.

Stimuli were made from press-on lettering on white 4×5 cards which were then photo-copied to provide two identical stimuli for the two stimulus channels.

Two types of stimuli were used: a 3×3 matrix of letters and a 3×3 matrix of black squares. The centers of the individual items were equidistant both horizontally and verti-cally. This resulted in a square display which subtended a visual angle of 6°26′ at a viewing distance of 30 in (1 in = 25·4 mm). The individual letters and squares subtended 1°12′. The fixation point, a period from the same set of lettering, subtended a visual angle of 7′. In the letter conditions the letters were assembled randomly with the added constraints that no meaningful words or abbreviations were included and no letter appeared more than once on a card. Three such sets of 25 stimuli each were constructed.

Procedure

Trials were run in blocks of 25 with stimulus duration, type of stimulus, channel of stimulation, and type of response held constant over each block. Subjects were told these conditions at the start of each block of trials. In the offset latency conditions, 10 catch trials were randomly given in each series. The duration of each catch trial was two and a half times as long as that of the related stimulus trials. Onset and offset latencies for a given condition were always run in successive blocks, using the same set of stimuli, but in a random order. The ordering of the other conditions was fully randomized. For all condi-tions subjects fixated with each eye a small dot in the center of a blank field equal in intensity to the stimulus field. The subject closed a telegraph key and was given a verbal warning signal followed approximately 1 s later by stimulus onset. At stimulus offset the fixation field was presented again. In conditions of monoptic stimulation all material was presented to the subject's dominant eye while the other eye was presented with a blank field of equal intensity and duration. In the dichoptic conditions each eye received 50% of the total stimulus duration with the offset of the stimulus to the first (non-dominant) eye synchronized with onset of the stimulus to the second (dominant) eye (Fig. 1). Subjects responded by releasing the telegraph key as rapidly as possible, and, in the letter condition, also reported as many letters as possible *after* the offset of the stimulus. In the square condition, no verbal report was required.

RTs less than 150 ms were discarded as anticipations and the very few RTs greater than 300 ms were discarded as errors. All such trials, constituting no more than 5% for each subject, were repeated later in the sequence to provide equal cell N's for data analysis. Approximately 1500 practice trials per subject were discarded before data collection was begun.

Three conditions of stimulation were run: monoptic letters, dichoptic letters, and mon-optic squares. All were run at stimulus durations of 100, 200, 400, 600, and 1000 ms. In addition, the monoptic squares were presented for durations of 1500, 2000, 2500, and 3000 ms. Order of stimulation and duration were randomized.

Results

Mean RT was calculated for each block of 25 trials. The mean RT to stimulus onset was subtracted from the corresponding mean offset RT and these differences were used in the analysis. The data are presented in Figures 2 and 3, which show
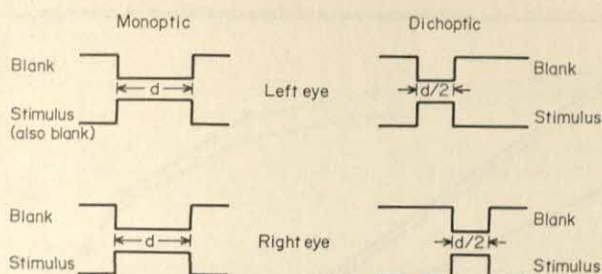
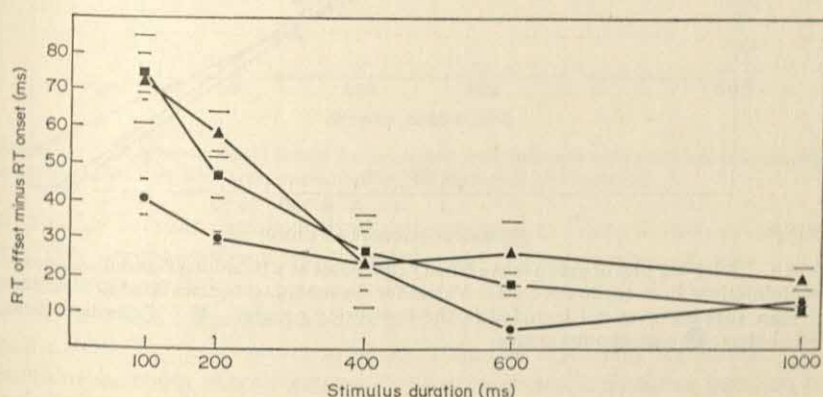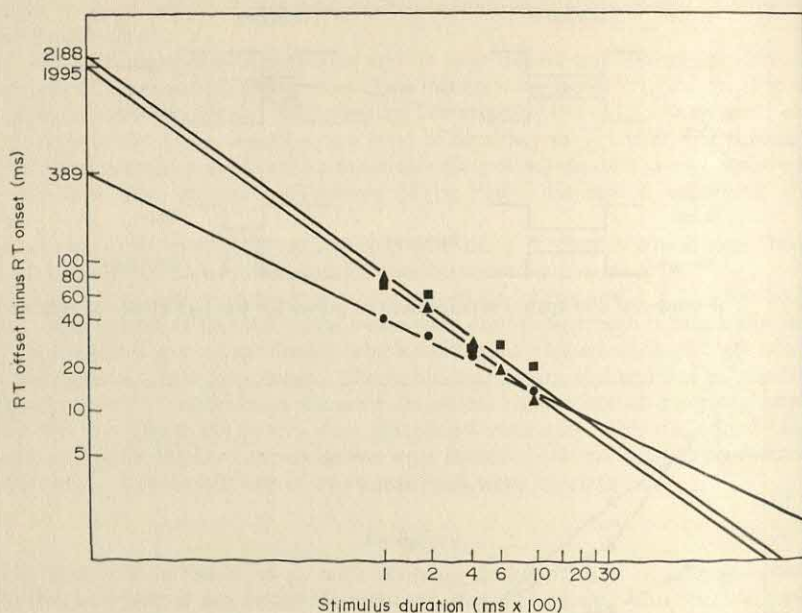FIGURE 1. The temporal and spatial arrangement of pulses for the two modes of presentation.

FIGURE 2. Group mean latency difference as a function of stimulus duration over a range of durations from 100 to 1000 ms. The brackets indicate the standard error of each mean. ● 9 monoptic squares; ▲ 9 dichoptic letters; ■ 9 monoptic letters.

TABLE I

Persistence values derived from the regression lines and observed persistences for each of five stimulus durations under each of three experimental conditions

| Stimulus duration (ms) | RT to offset minus RT to onset | | | | | |
| | 9 letters dichoptic | | 9 letters monoptic | | 9 squares monoptic | |
	Der.	Obs.	Der.	Obs.	Der.	Obs.
100	80·40	72·53	72·41	74·21	41·15	41·32
200	47·25	58·23	43·89	46·78	29·32	30·40
400	29·32	23·71	26·73	26·49	20·91	23·51
600	21·88	26·16	19·95	17·78	17·10	5·72
1000	15·12	18·96	13·77	10·95	13·32	12·73
mean error		6·52		1·99		3·16
1500						8·73
2000						4·44
2500						−0·97
3000						1·61

FIGURE 3. A log-log plot of group mean latency difference as a function of stimulus duration over a range of durations from 100 to 1000 ms. Values for the monoptic squares condition for durations greater than 1000 ms were not included in the line-fitting process. ■ 9 dichoptic letters; ▲ 9 monoptic letters; ● 9 monoptic squares.

mean differences for the group of three subjects for all conditions as functions of stimulus duration. Figure 3 shows the same data plotted on log-log co-ordinates. Equations for the least squares regression lines for stimulus durations from 100 to 1000 ms are power functions having slopes of -0.72, -0.72, and -0.49 and y-intercepts of 2188, 1995 and 389 for the dichoptic letters, monoptic letters, and monoptic squares respectively. The points for exposures greater than 1000 ms (squares only) are shown only in Table I and were not included in the analysis of variance (Table II). A repeated measures analysis of variance (Table II) indicates

TABLE II

Repeated measures analysis of variance on persistence values for two types of stimuli and five stimulus durations

	SS	df	ms	F	P
A (stimulus duration)	6539·53	4	1634·88	6·30	<0·05
B (stimulus material)	619·26	1	619·26	2·55	>0·05
Material × duration interaction	391·84	4	97·96	0·17	>0·05
Subjects	977·79	2	488·90		
Subjects × duration	2076·18	8	259·52		
Subjects × material	485·02	2	242·51		
Interaction × subjects	4588·98	8	573·62		
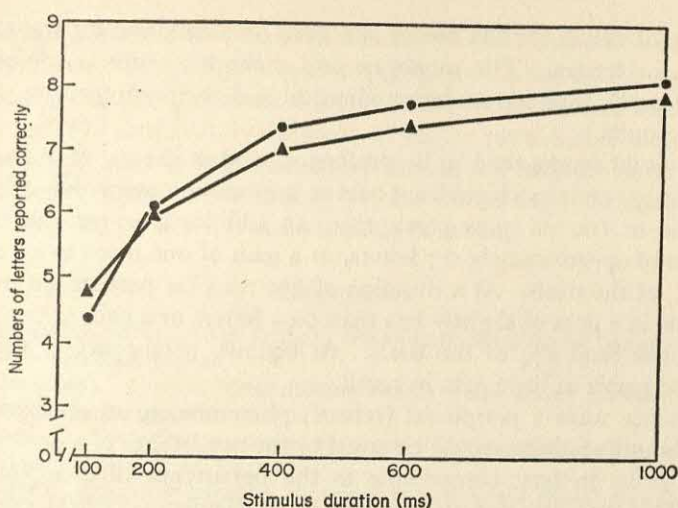
FIGURE 4. Group mean recall scores for monoptic and dichoptic letter conditions as functions of stimulus duration. ▲ Monoptic presentation; ● Dichoptic presentation.

that stimulus duration significantly affects persistence. Type of material (squares or letters) has a non-significant effect and the interaction is non-significant.

For all three conditions, exposure duration has a marked effect both on persistence of the visual stimulus and on the recall scores for the two letter conditions. Figure 4 shows the recall scores to be an increasing negatively accelerated function of stimulus duration. Above 400 ms, further increases in stimulus duration result, cumulatively, in an increase of slightly less than one letter for an *additional* stimulus duration of 600 ms. As performance at 400 ms approaches total recall, the increment of less than one letter per additional 600 ms is not very revealing.

Discussion

The estimate of persistence obtained in this study is slightly conservative, in that RT to the offset of a stimulus is consistently shorter than RT to the onset of the same stimulus (Jenkins, 1926), the difference being of the order of 5 ms. The storage times obtained from RT measurements are substantially less than those obtained by the less direct measures which were discussed earlier. The largest value obtained for the presentation of the matrix of squares is on the order of 40 ms which is approximately one-sixth the duration estimated for the Sperling store. This contradicts Neisser's (1967) statement that STVS is indistinguishable from the physical stimulus. If we accept the existence of a short-term store of approximately 250 ms in duration, such a store, while perhaps phenomenally similar, cannot be an indistinguishable phenomenal persistence of the visual stimulus. However, persistence of the visual stimulus does appear to account for some part of post-stimulus information storage.

The relation between persistence and stimulus duration (Fig. 3) is a power function of the form $P = kD^n$. Persistence is equal to the product of a constant (the y-intercept) and stimulus duration raised to the nth power (the slope of the line). As predicted, persistence is inversely related to stimulus duration. Two

different sets of values for the power law were obtained, one for the squares and another for the letters. The monoptic and dichoptic letter conditions did not differ significantly, their slopes being identical and their y-intercepts differing by only 0·04 log units.

All brief visual events tend to be prolonged, and at shorter stimulus durations persistence may result in a significant gain in information extraction. For example, at a duration of 100 ms (plus persistence) an additional 40 ms will result in an average gain of approximately 0·7 letters, or a gain of one more letter on approximately 70% of the trials. At a duration of 600 ms plus persistence an additional 40 ms results in a gain of slightly less than 0·05 letters or a gain of one more letter on slightly less than 5% of the trials. At 600 ms, persistence of approximately 15 ms would result in little gain in recall.

If persistence were a peripheral (retinal) phenomenon, the persistence of the 200 ms dichoptic stimulus would be equal to the persistence of a 100 ms monoptic stimulus. It is, in fact, comparable to the persistence of a 200 ms monoptic stimulus. This indicates that the persistence mechanism is central, operating after confluence of the inputs from the two eyes.

A non-significant trend toward greater persistence of the letters than of the squares was observed. If further work shows this to be real, it might be due to a division of attention caused by the added information processing task in the letter condition. It is unlikely that all delays in offset RTs for both squares and letters are due to division of attention or to central intermittency rather than to persistence of the visual stimulus. In view of evidence from numerous intermittency studies (Bertelson, 1967; Smith, 1967), delays in RT due only to central intermittency are not accurately described by power functions such as those obtained in this study. In the present procedure, all trials were blocked in terms of stimulus duration and, therefore, ISI. Several studies (e.g. Borger, 1963) have shown that the consequent reduction in time uncertainty eliminates or greatly reduces intermittency effects. Further, Borger (1963) found no delay in RT_2 when subjects were not required to respond to S_1. This is the exact paradigm used in the square condition. In the letter conditions an encoding response was required to S_1 followed by a response to S_2. Thus, small delays of the type observed in this study are to be expected in the letter conditions but not in the square condition. Finally, the curve that describes the data for the square condition was not of the form expected from the intermittency paradigm, but a straight line on a double log plot which approaches zero only when ISI (stimulus duration) is greater than 1000 ms. At this point the persistence is very small but continues to decrease as far as 3000 ms (the longest stimulus duration used in this study). The continuing decrease to 3000 ms indicates that the delays are not due to stimulus onset functioning as a warning signal for the offset latency response. Were this the case, offset RTs would begin to increase again at approximately 1500 ms (Woodworth and Schlosberg, 1954).

Conclusion

Visual persistence is inversely related to stimulus duration. For durations from 100 to 1000 ms this relationship is described by a power law $P = kD^n$. The value

of n obtained in this study is $-0\cdot49$ which makes a close approximation of the law state that persistence is equal to a constant (389) multiplied by the square root of the reciprocal of stimulus duration. Monoptic and dichoptic methods of presentation do not result in significantly different values for persistence. The addition of an information processing requirement results in an addition of a component presumably due to central intermittency but does not change the basic form of the relationship between persistence and stimulus duration.

References

AVERBACH, M. and CORIELL, A. S. (1961). Short-term memory in vision. *Bell System Technical Journal*, **40**, 309–28.

BERTELSON, P. (1967). Central intermittency twenty years later. *Quarterly Journal of Experimental Psychology*, **18**, 153–63.

BORGER, R. (1963). The refractory period and serial choice reactions. *Quarterly Journal of Experimental Psychology*, **15**, 1–12.

EFRON, R. (1970). Effects of stimulus duration on perceptual onset and offset latencies. *Perception and Psychophysics*, **8**, 231–4.

HABER, R. N. and STANDING, L. G. (1969a). Direct measures of short-term visual storage. *Quarterly Journal of Experimental Psychology*, **21**, 43–56.

HABER, R. N. and STANDING, L. G. (1969b). Direct estimates of apparent duration of a flash followed by visual noise. Unpublished paper cited in HABER, R. N. (1969). Visual persistence and the processing of visual stimulation. Paper presented at the International Congress of Psychology Convention, London, 1969.

JENKINS, T. N. (1926). Facilitation and inhibition. *Archives of Psychology*, No. 86.

KEELE, S. W. and CHASE, W. G. (1967). Short-term visual storage. *Perception and Psychophysics*, **2**, 383–6.

NEISSER, U. (1967). *Cognitive Psychology*. New York: Appleton-Century-Crofts.

SMITH, M. (1967). Theories of the psychological refractory period. *Psychological Bulletin*, **67**, 202–13.

SPERLING, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, **74**, No. 11 (Whole No. 498).

WOODWORTH, R. S. and SCHLOSBERG, H. (1954). *Experimental Psychology*. New York: Holt, Rinehart & Winston.

THE FOUR-CARD PROBLEM AND THE GENERALITY OF FORMAL REASONING

E. A. LUNZER, C. HARRISON AND M. DAVEY

School of Education, University of Nottingham

Three experiments were designed to investigate the failure of intelligent adults to solve an apparently simple problem of formal reasoning devised by Wason. Both the mode of presentation and the type of material reduced the difficulty of the problem, while retaining its essential form. However, success on the original problem remained at a low level, even when subjects had attempted an easier version and had been given an explanation.

These results enable one to reject a "strong" formulation of Piaget's theory of formal reasoning. A "weaker" formulation is suggested as a basis for further research.

Introduction

In a recent series of papers, Wason (1968, 1969*a*, *b*) has brought out the unexpected difficulty of a very simple problem in implication. The subject is faced with four cards showing respectively a red triangle, a red circle, a blue triangle and a blue circle. These are taken from a set of eight, all of which bear a triangle on one face and a circle on the other. This is known to the subject, the colours on the hidden faces being the only unknowns. The problem is to say which of the four cards must be turned over to establish whether the following proposition is true or false: "*Every card which has a red triangle on one side has a blue circle on the other side*". Less than 5% of subjects at university level give the correct reply "*red triangle, red circle*" (the only two cards which might feature a red triangle on one side and a not-blue circle on the other, so falsifying the proposition). The reply is most frequently "*only the red triangle*" or "*red triangle, blue circle*" although the truth-value of the proposition is unaltered when the triangle on the reverse of the blue circle is also blue.

Previous research by Wason (1968) has shown that the difficulty of the problem is not due to the form of the test question. The wording "Every card which has a triangle on one side . . ." might, in ordinary speech (though not in logic) be taken to imply that at least two of the cards do in fact have a red triangle and a blue circle. The wording "If a card has a red triangle on one side, it has a blue circle on the other side" would appear less misleading. No difference was found. This was confirmed by the authors in a preliminary study in which 42 subjects were given the problem in the form "*Every card . . .*" and 42 in the conditional form. All the subjects were drawn from the 5th and 6th forms of a selective grammar school. Each group produced exactly three correct answers.

Striking as they are, these results are in line with an accumulating body of evidence on the treatment of hypotheses in human reasoning. Having once formed a hypothesis on the basis of preliminary evidence, the subject tends to stay with it

until it is disconfirmed. In seeking further evidence he selects confirmatory instances rather than genuine tests (Wason, 1960). More specifically, where the hypothesis predicts a positive outcome under given conditions, he does not actively test for a negative result outside those conditions. There is thus a failure to consider alternative hypotheses which is reminiscent of "successive scanning", the most primitive strategy found in Bruner, Goodnow and Austin's classic work on concept attainment (1956).

However, the evidence derived from the four card problem has a special interest because it arises in the context of a closed system. Thus, even where the situation appears to favour an application of the formal structure of the propositional calculus, the human reasoner fails to apply it. As Wason remarks, "such a finding is difficult to reconcile with Piaget's theory of formal operations" (1969b). There is no combinatorial analysis of the variables, said by Inhelder and Piaget (1958) to be characteristic of post-adolescent reasoning.

The following experiments are part of a programme which has three aims in view: (1) to elucidate the source of difficulty in the four card problem; (2) to examine the effectiveness of various procedures designed to alert the subject to its underlying logical structure; and (3) to establish whether there are age-related differences with respect to the effectiveness of such procedures. It is clear that the answer to the third of these problems may point the way to an integration of these findings with those of Piagetian enquiries into adolescent thinking (Lunzer, 1968; Peel, 1971). However, in the present experiments, attention is confined to the first two problems, these being a necessary preliminary.

Experiment I

Piaget maintains that post-adolescent thinkers must have elaborated the propositional calculus for themselves, since they spontaneously apply it in Inhelder's experiments. Given this assumption, their failure in the four card experiment might be attributed to one or both of two causes. The first relates to the content of the problem, the second to the form of presentation.

The proposition to be tested has the form "p implies q", where $p = $ "*the triangle is red*", not-$p = $ "*the triangle is blue*", $q = $ "*the circle is blue*" and not-$q = $ "*the circle is red*". Each of these atomic propositions corresponds to an arbitrary association between independent properties of a formal diagram. One may therefore suppose that combining them in an implication will impose a severe strain on short-term memory, and so impair the efficiency of reasoning. If so, the difficulty should be greatly reduced by substituting an alternative content which would preserve the form of the problem, but would figure more realistic associations. Such are the cards shown in the lower half of Figure 1. One side shows a lorry in red or yellow, while the other side shows the inside of the lorry which is either empty or full of coal.[1]

A second possibility relates to the surface simplicity of the problem. The subject is misled into thinking it can be solved by simply turning cards over to see which

[1] We are indebted to Mr M. Abbott for the suggested variation in material. A somewhat similar experiment by Wason and Shapiro (1971) is referred to in the discussion but was not known to us when this research was undertaken.

Shapes complete

Shapes reduced

Lorries complete

Lorries reduced

Red	//////	Blue		Yellow	▓▓▓
Empty		Full of coal	██		
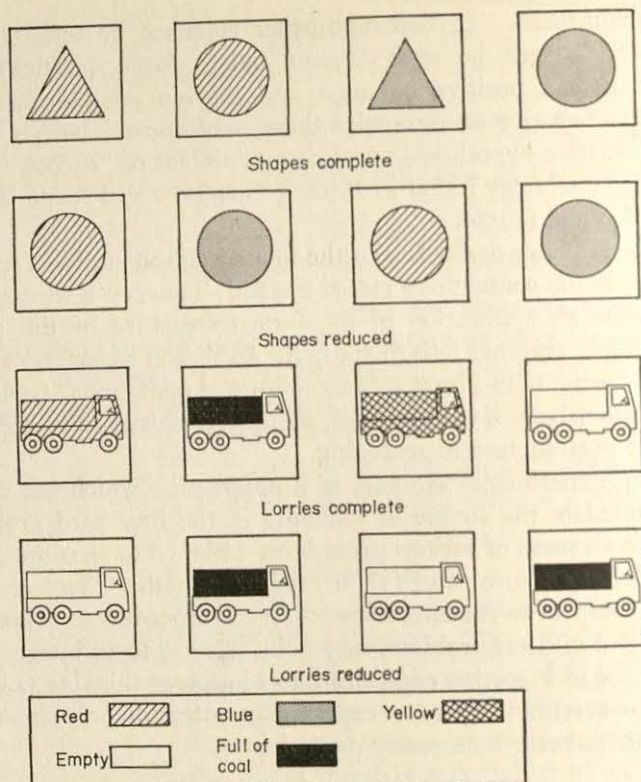
FIGURE 1. Materials and modes of presentation

ones verify the proposition, instead of being led to consider its consequences, and to deduce the importance of the cards which might falsify it. Wason and Johnson-Laird (1970) have shown that once a selection strategy has been determined it is difficult to invalidate it by a subsequent evaluation process, because the selection derives from an auto-instruction to which the subject lacks access. Given the *complete* presentation with four different cards (Fig. 1), the subject will naturally want to turn over the red triangle. This card is indeed relevant, but for the opposite reason to that which determined its selection, viz. that it could falsify the proposition, not that it could verify it. The presence of this card combined with its logical necessity has the effect of reinforcing the incorrect auto-instruction. The *reduced* presentation in Figure 1 should alert the subject to the real nature of the problem. The formal element remains unchanged because he must still decide which *pair* of cards is relevant to the truth value of the proposition.

It was further hypothesized, still on the assumption that Piaget's view is substantially correct, that once alerted to the structure of the problem either by the realistic material or by the reduced presentation, the subject would readily transfer a correct solution strategy when faced with the problem in the more difficult form. It was therefore decided to test the effectiveness of the mere presentation of an easier related problem, compared with an explanation similar to that used by Wason (1969a).

Method

The subjects were 64 sixth formers in a selective boys' grammar school, all of whom had at least five 'O' level G.C.E. passes. These were divided into two main groups of 32, each containing a similar range of academic ability and a similar mixture of arts and science students. All subjects were given two problems to answer. In the case of one of these main groups the correct answer to the first test was given and fully explained before proceeding to the second. The other group were given no feedback or instruction between the two tests.

Within each of these two main groups, half were set the first problem using the realistic material and the second with the formal shapes, and half were set the problems in reverse order. Finally, there were four combinations of presentation. Effectively, the 64 subjects were therefore allocated to one of 16 groups, four subjects in each group, i.e.

Group 1	Lorries reduced + shapes complete.
Group 2	Lorries complete + shapes reduced.
Group 3	Lorries complete + shapes complete.
Group 4	Lorries complete + shapes complete.
Groups 5–8	As above, order of problems reversed.
Groups 9–12	As groups 1–4, with explanation intervening.
Groups 13–16	As 9–12, order of problems reversed.

Procedure

(1) *Training.* Before each problem was presented, subjects were made familiar with a set of eight cards, out of which four would be used in the problem. The eight cards feature two instances of each type of element in the domain e.g. two red lorries, full, two red lorries, empty; two yellow lorries, full, two yellow lorries, empty.

The set in use was placed on the table in random array and it was pointed out that there was no correlation between colour and fullness (lorries), or between colour of triangle and colour of circle (shapes).

(2) *Problem.* The four test cards were arranged in a row before being revealed to the subject. The test sentence was now shown and the test question presented in written and spoken form: "Which of the cards do you need to turn over in order to find out whether the sentence in front of you is true or false?" If a subject chose only one card he was asked "Only that one?" in case he had interpreted "Which of the cards" as meaning "Which one of the cards".

(3) *Explanation* (Groups 9–16). When the subject declared himself satisfied with his choice, the experimenter drew his attention to the first card in the row (regardless of his choice), asking first, what could be on its other side (correction when necessary), and second, what effect each alternative would have on the truth value of the test proposition. If one of the alternatives made the proposition false, the card was pushed forward as requiring examination, otherwise it was pushed back. The cards were not turned over. The same procedure was then applied to cards 2–4. Finally, the subject was asked which of the four cards needed to be turned over and to explain why. If the explanation at this point was inadequate in respect of any card, the teaching steps were repeated in respect of that card. Total time for the teaching varied from 5 to 10 min.

Results

Table I shows the frequency of the different types of solution as a function of the main variables. It is evident that the proportion of successes far exceeds those found in previous studies, indicating the facilitating effects of modified presentation. Thus, for the first problem, only three subjects out of 16 select the falsifying not-q card when the problem appears in the original form, while 13 choose the not-q pair in the reduced lorries condition.

TABLE I

The frequencies of three types of selection (1, $\frac{1}{2}$, 0) as a function of presentation, material, order of problem, and previous explanation

(a) *All groups*

Presentation	Material	First problem O	First problem $\frac{1}{2}$	First problem I	Second problem O	Second problem $\frac{1}{2}$	Second problem I
reduced	lorries	3	4	9	2	3	11
reduced	shapes	5	8	3	3	1	12
complete	lorries	7	5	4	5	8	3
complete	shapes	13	1	2	10	2	4
Total ($N = 64$)		28	18	18	20	14	30

(b) *Groups 1–9.* No explanation given

Presentation	Material	First problem O	First problem $\frac{1}{2}$	First problem I	Second problem O	Second problem $\frac{1}{2}$	Second problem I
reduced	lorries	2	1	5	0	2	6
reduced	shapes	3	3	2	1	1	6
complete	lorries	1	4	3	2	6	0
complete	shapes	6	1	1	5	2	1
Total ($N = 32$)		12	9	11	8	11	13

(c) *Groups 9–16.* Explanation after Problem 1

Presentation	Material	First problem O	First problem $\frac{1}{2}$	First problem I	Second problem O	Second problem $\frac{1}{2}$	Second problem I
reduced	lorries	1	3	4	2	1	5
reduced	shapes	2	5	1	2	0	6
complete	lorries	6	1	1	3	2	3
complete	shapes	7	0	1	5	0	3
Total ($N = 32$)		16	9	7	12	3	17

The three main types of solution are given the following symbols:

$1 = p$, not-q (complete); not-q, not-q (reduced);
$\frac{1}{2} = p$, q, not-q (complete); all cards (reduced);
$0 = $ others.

Taking the first presentation of the problem, and referring to Table 2, the overall value of χ^2 for the first block of data in Table I is 22·0 ($P < 0\cdot001$, one-tail test). Johnson-Laird and Wason (1970a) indicate how the distribution of replies to the four card problem suggests the intervention of two successive insights. Insight 1 is the realization that any card which could falsify the proposition should be selected. Insight 2 is the realization that only these cards should be selected. To test for Insight 1 (clearly more frequent in our data than when the problem is given in the original form), the table may be partitioned vertically by pooling the

<div align="center">TABLE II</div>

<div align="center">*Partitioning of Chi-square*</div>

(a) *First problem*

	df	value of χ^2	P
Overall	6	22·0	<0·001
insight 1 (pooling $\frac{1}{2}$ and 1 columns)	3	14·22	<0·01
reduced vs. complete (within insight 1)	1	7·86	<0·05
material within reduced (,, ,, 1)	1	0·17	NS
material within complete (,, ,, 1)	1	3·33	<0·05
insight 2 ($\frac{1}{2}$ and 1 columns only)	3	4·64	NS
reduced vs. complete (within insight 2)	1	0	NS
material within reduced (,, ,, 2)	1	2·68	NS
material within complete (,, ,, 2)	1	Fisher's exact	P=0·409 (NS)

(b) *Second problem*

	df	value of χ^2	P
overall	6	24·55	<0·001
insight 1 (pooling $\frac{1}{2}$ and 1 columns)	3	9·89	<0·01
reduced vs. complete (within insight 1)	1	4·16	<0·05
material within reduced (,, ,, 1)	1	<1	NS
material within complete (,, ,, 1)	1	2·76	<0·05
insight 2 ($\frac{1}{2}$ and 1 columns only)	3	12·68	<0·01
reduced vs. complete (within insight 2)	1	7·38	<0·01
material within reduced (,, ,, 2)	1	Fisher's exact	P=0·270 (NS)
material within complete (,, ,, 2)	1	Fisher's exact	P=0·127 (NS)

Yates' correction used throughout where $df = 1$

last two columns ($\chi^2 = 14·22$, 3 df, $P < 0·01$). The presence of Insight 2 (given Insight 1) may now be tested by computing the partial χ^2 for the last two columns only (previously pooled), i.e. (4, 9), (8,3), (5,4), (1,2). This has the value of 4·64 (3 df) which is not significant.

Each of the two partial χ^2 values may be further partitioned to yield three 2 × 2 tables, viz. reduced vs. complete, shapes vs. lorries within reduced, and shapes vs. lorries within complete. The three χ^2 values (1 df with Yates' correction) for Insight 1 are 7·86 ($P < 0·01$), 0·17 (NS), and 3·33 ($P < 0·05$), respectively. Hence, Insight 1 (appreciation of the relevance of the not-q card(s)) is favoured by the reduced array, and is favoured by the realistic material when the array is complete. However, when the reduced array is shown, the additional facilitation of content is marginal. For Insight 2 (irrelevance of the q card(s)), there is no difference between the reduced and complete arrays. The only significant effect is within the reduced presentation, when the effect of realistic material yields a χ^2 of 2·68 ($P < 0·05$, one-tail test). Thus it appears that both facilitating conditions are essential to produce Insight 2 with appreciable frequency.

For the second problem, the overall χ^2 value is 24·55 ($P < 0·001$). The first partitioning yields a value of 4·16 ($P < 0·05$) for Insight 1 and 12·68 ($P < 0·01$) for Insight 2. Thus, a completely logical solution is favoured by the changes in

conditions, given previous experience of an analogous problem. Inspection of Table II indicates that, for Insight 1, presentation and material interact very much as they did for the first problem. However, Insight 2 is clearly very much more frequent under reduced presentation.

It was an inherent feature of the design that some of the subjects went from an easier problem to a more difficult one, some from a difficult to an easier one, and some from a problem to one which was easier in one respect but harder in another. Taken overall, however, the incidence of problem difficulty was the same for both occasions. Moreover, since it was hypothesized that the experience of an easier first problem would favour solution of one more difficult, a high incidence of change from failure to success would be predicted.

To treat the two insights separately, within the same analysis, would have resulted in very low frequencies in the critical cells. It was therefore decided to look for changes under two criteria: total failure versus partial or complete success, and complete success vurses other solutions. The first criterion corresponds to a *heuristic* mode of verification in which the inductive mode is retained alongside the deductive, i.e. confirming instances are deemed to strengthen a hypothesis. The second is a strict *logical* criterion based on the propositional calculus which asserts that an implication is false if and only if the antecedent is true and the consequent is false.

Table III shows the incidence of change based on the first heuristic criterion (Insight 1 or Insights 1 and 2) and on the strict logical criterion (Insight 2). Unlike the analyses in Table II, there is no strict partitioning in that the two analyses are not independent (individuals listed as + are pooled with success on the heuristic criterion, and re-appear as failures on the logical criterion).

The data are not easy to interpret. Whatever the group, and whatever the criterion, gains exceed losses. On the other hand, the margin is rarely large. On the heuristic criterion, none of the χ^2 values are significant, and one may conclude.

TABLE III

Incidence of change from first to second problem

	$--$	$-+$	$+-$	$++$	χ^2	(McNemar)
Logical criterion						
Groups 1–8 (no explanation)	15	6	4	7	0·225	NS
Groups 9–16 (explanation given)	14	11	1	6	7·52	(P < 0·005)
All groups	29	17	5	13	6·01	
Heuristic criterion						
Groups 1–8 (no explanation)	2	10	6	14	0·766	NS
Groups 9–16 (explanation given)	7	9	5	11	0·875	NS
All groups	9	19	11	25	1·875	

$++$ = Positive score on both problems.
$+-$ = Positive score on first problem and negative score on second.
$-+$ = Negative score on first problem and positive score on second.
$--$ = Negative score on both problems.

that the number of subjects who regress to a completely "intuitive" (incorrect) mode of response, when presented with a difficult problem following an easier one, is nearly as great as the number who improve when the second problem is easier.

The logical criterion produces significance, restricted to those groups who were given the explanation procedure. However, since complete success was relatively rare on the first presentation of the problem, the incidence of negative change is greatly reduced, and this result must be interpreted with caution. Positive change would appear to depend on direct teaching as against mere experience of an analogous problem.

Discussion

The results of the experiment support two hypotheses concerning the difficulty of the four card problem: both the realistic material and the reduced presentation lead to significant increases in correct solutions. On the other hand, they do not confirm the third hypothesis, since experience by itself does not produce significant improvement in a new problem.

However, the complexity of the design precluded any sharp analysis of learning effects. In particular, only 16 subjects were given the critical problem second, i.e. geometrical shapes with complete presentation, only eight of these had been given an explanation, and the type of problem preceding was a further variable.

It was therefore decided to consider in the next experiment the effects of presentation and material in producing transfer to the "shapes complete" problem.

Experiment II

Method

The subjects were 60 sixth-form girls, all of whom satisfied the same criteria as the boys in Experiment I. These were divided into four representative groups roughly equated in terms of science/arts bias and academic ability. Groups 1–3 were of 16 girls, while Group 4 contained 12.

All of the subjects in Groups 1–3 were given a first problem under the same conditions as those in Experiment I. This was followed in all cases by the explanation procedure previously outlined. This, in turn, was followed for all the groups by an identical "complete shapes" second problem: red triangle, blue triangle, red circle, blue circle.

Thus, the only variable was the first problem. For Group 1, this was "lorries complete", i.e. the material was assumed to be facilitating, while presentation was held constant. For Group 2, it was "shapes reduced" i.e. the material was assumed to be facilitating, while presentation was held constant. For Group 3, it was a "shapes complete" problem, with the material: green diamond, green square, yellow diamond, yellow square, and the test proposition: "*Every card which has a green diamond on one side has a yellow square on the other*".

Group 4 were shown the same material as Group 3 but proceeded directly to explanation after the material had been presented, i.e. they were not asked to solve the problem until after the explanation had been given. Thus, unlike the girls in Group 3, none of those in Group 4 had previously committed herself to a (probably) incorrect solution.

Results

The results shown in Table IV are striking and unexpected. Group 3, to whom the solution was explained in the context of a problem which was similar in all respects to the second problem, produced more than 50% correct solutions to the second problem. Groups 1 and 2, who received the same explanation, but in the context of an easier problem, produced 75% "intuitive" (incorrect) solutions. Even more surprising is the failure of Group 4.

Heuristic solutions were relatively infrequent in all groups. The following comparisons are therefore based on pooling the "1/2" and "1" columns. The girls in this experiment were less successful than the corresponding groups of boys in Experiment I. Sixteen girls and sixteen boys attempted the complete lorries problem, but only five girls were successful as compared with nine boys. Sixteen of each sex attempted the shapes reduced, and only three girls were successful as against eleven boys. Although the first difference is not significant ($\chi^2 = 1 \cdot 143$),

TABLE IV

Frequency of partial (heuristic) and complete (logical) success

Group		Problem 1			Problem 2 (shapes complete)		
		0	½	1	0	½	1
Presentation and material							
1	lorries complete	11	4	1	13	0	3
2	shapes reduced	13	1	2	11	1	4
3	shapes complete	14	1	1	7	0	9
4	explanation only	—	—	—	9	0	3

the second is significant ($\chi^2 = 6 \cdot 22$, 1 *df*, $P < 0 \cdot 02$, two-tailed test). No appreciable sex difference was apparent for the shapes complete problem when presented first.

Despite the striking failure of Groups 1, 2, and 4 in the criterion problem, comparisons with Group 3 do not reach significance. For Group 1 vs. Group 3, $\chi^2 = 3 \cdot 33$, for Goup 2 vs. Group 3, $\chi^2 = 1 \cdot 14$ and for Group 4 vs. Group 3, $\chi^2 = 1 \cdot 26$ (where a value of $3 \cdot 84$ would be required for significance at the $0 \cdot 05$ level).

Discussion

There is one qualitative observation which needs to be brought out. Subjects who failed in Problem 1 in the complete presentation nearly always offered the solution "*p, q*" (e.g. "red lorry, full lorry"). The same was true of the boys in Experiment I. The reply "*p* only", found by Wason (1969*a*), was infrequent. In the present experiment, the solutions to the criterion problem, offered by subjects

in Group 2 who failed on the first problem, did not show a predominance of the "p, q" solution ("red triangle, blue circle"). The actual distribution was:

p, q	4	
p, not-q	4	(= Insight 2)
p, q, not-q	1	(= Insight 1)
not-p, not-q	5	
all	1	
p, not-p	1	

Thus, no less than five of the eleven failing subjects offered the unexpected solution "blue triangle, red cicle" and one offered "blue triangle, red triangle".

Group 2 had attempted the reduced shapes problem and the correct solution not-q not-q had been explained. The reduced array consists of 4 figures. The "intuitive" (incorrect) solution is to select the two cards which could verify the proposition and reject the others. The correct solution is to select only these others, since they could falsify the proposition. Similarly, in the complete presentation, the "intuitive" (incorrect) solution is "red triangle, blue circle", because only these cards could verify the proposition (and are mentioned in it). The error "blue triangle, red circle" suggests that these subjects had failed to understand the explanation, but had retained the solution, together with a simplified algorithm which might sustain a limited transfer, having the form "reverse the values of the obvious (intuitive) solution". An analogous algorithm: "Retain the value of the antecedent in the proposition; reverse the value of the consequent", would produce correct solutions to the second problem in Group 3. It follows that one cannot safely infer that all of the nine subjects in the group who gave correct solutions did so because they had fully mastered the previous explanation.

Experiments I and II differ from those previously reported in the literature in that the subjects were sixth-formers and not university students. Hence, one cannot rule out the possibility that the relative failure to transfer from an easier problem to the more difficult "shapes complete" problem is due to lack of selectivity in the sample, as compared with the subjects of previous studies. The next experiment was carried out with university graduates as subjects. It was designed to examine the effects of explanation and the effects of transfer from an easier problem.

Experiment III

Method

Three problems were presented to the subjects as a class exercise, to be solved individually by each member of the group. Three types of problem were set in the following order: (1) the "lorries reduced" problem; (2) the "shapes complete" problem; (3) an alternative version of the "shapes complete" problem, using the sentence "*Every card which has a black square on one side has a white diamond on the other*". In between the second and third problem, the correct solution to the first problem only was given and fully explained.

The subjects were 31 graduate students enrolled in a post-graduate certificate course in education. There were approximately equal numbers of males and females, and of arts and science graduates.

Procedure

Problem 1. The four lorries were drawn on the board and labelled. The test proposition was written down and read out. The students were carefully told the nature of the problem, i.e., to test the proposition with reference only to the four cards illustrated. They were also told the constraints, i.e., that only one side of each card was shown, that each card had another side, that the only unknown about the other side was whether it was "red or yellow". Every care was taken to ensure that each member of the group fully understood what was required, and questions were answered to clear up misunderstandings. When the task was clear to all members of the group, they were told to work out the solution, taking their time, and to write it on a prepared answer sheet. Approximately 5 min were given for the solution.

Problem 2. Immediately all subjects in the group were satisfied with their solution to Problem 1, the group was shown Problem 2, the complete shapes problem, following exactly the same procedure. Once again, questions about the constraints were answered, and the subjects were told to work out the solution and enter it on the sheet provided.

Explanation. Subjects were instructed not to alter anything they had written. The experimenter then carefully went through the diagrams used for Problem 1. Taking each diagram in turn, he elicited from the group "what could be on the other side", what could be said of the proposition in each case and whether there was any point in turning the diagram over. This allowed the experimenter to bring out the fact that only the two empty lorries needed to be examined. Questions were answered until every member of the group declared himself satisfied with the explanation.

Problem 3. This was given immediately after the explanation of Problem 1, using the same procedure as before.

Results

The results of the experiment are given in Table V. Thirteen subjects (37%) showed at least Insight 1 (heuristic criterion) in Problem 2. The indications from previous research with comparable subjects are that less than 10% would be

TABLE V

The frequency of logically correct (1), heuristically correct (1/2), and incorrect solutions (0)

	0	$\frac{1}{2}$	1
Problem 1 (reduced lorries)	9	11	11
Problem 2 (complete shapes)	18	10	3
Problem 3 (complete shapes, following explanation of problem 1)	17	7	7

expected to succeed even on the heuristic criterion. Thus, it is likely that the previous experience of the "reduced lorries" problem produced a transfer of insight in an appreciable proportion of subjects. In addition, two subjects were partially successful in Problem 2 despite their failure in Problem 1. However, more than half the subjects failed Problem 3 even after this initial experience. Five failed Problem 2, having attained the heuristic solution to Problem 1, and six having attained the correct solution.

Problem 3 was in all essentials a replica of Problem 2, but with the explanation intervening. Of the 18 subjects who failed on Problem 2, no less than 14 failed on Problem 3. On the other hand, four subjects moved from the heuristic to the

correct solution. Thus, the effect of explanation when added to experience, appears to be largely confined to making an existing understanding more precise. When the experience of Problem 1 failed to produce at least a partial insight, its explanation tended to be ineffective. Finally, there were no reliable differences in performance between males and females, or between Arts and Science graduates.

Discussion

Disregarding minor variations due to sampling error, the results of the three experiments are consistent and may be summarized as follows:

(1) Success in the original four-card problem, when not preceded by a similar problem, is rare. The findings of the present study agree with previous findings.

(2) The substitution of realistic material leads to a significant increase in the proportion of successes.

(3) The use of a reduced presentation produces a similar increase.

(4) When both effects are combined ("lorries reduced"), the incidence of failure is well below 50% (19% in Experiment I, 29% in Experiment III).

(5) Previous experience of an easier problem produces some increase in the solution of a criterion problem, but the rate of failure is still greater than 50%.

(6) Even experience combined with explanation does not produce a 50% success rate. Explanation does, however, tend to produce Insight 2 (irrelevance of a potential confirming card) provided the subject has already attained Insight 1 (relevance of a potentially falsifying card).

(7) Success on the criterion problem following experience and explanation of a virtually identical problem produces a success rate above 50%. However this condition cannot be taken as a satisfactory proof of genuine understanding.

Because these experiments are part of a series which is not yet complete, discussion will be brief. (1) adds nothing to existing knowledge. (2), (3), and (4) provide confirmation of two of the original hypotheses. The abstract formal content adds to the difficulty of the problem, and the problem can be facilitated by alerting the subject to the source of the difficulty. The effect of realistic content (2) is paralleled by the recent finding of Wason and Shapiro (1971), who compared thematic material with abstract material and found a significant advantage to the former. The effect of the reduced presentation is similar, but not identical, to the result of an experiment by Johnson-Laird and Wason (1970b), in which choices were required from two categories of objects, corresponding to "q" and "not-q".

All these findings show that the quality of reasoning is greatly affected by content, whether abstract or realistic, and this is inconsistent with a pure Piagetian account, e.g. "when the subject becomes capable of reasoning in a hypothetico-deductive manner, ... he relies on the necessary validity of an inference (vi formae) as opposed to the agreement of conclusions with experience." (Piaget, 1950, p. 148). The

greater difficulty of abstract material is borne out in the work of Roberge and Paulus (1971).

The very limited efficiency of experience and explanation in reducing failure on the four-card problem constitutes a totally unexpected set of findings, i.e. points (5), (6), and (7) above. In Wason's (1969a) study, 28 of 32 subjects were able to explain the problem after demonstration in which the cards were turned and 16 after verbal explanation only. In the Wason and Johnson-Laird study (1970b), 21 out of 34 learned by verbal explanation. The present findings suggest that the criteria of such studies were insufficiently stringent. In Experiments I and II all the subjects who received an explanation of the problem appeared to understand it. But most of them failed to apply this supposed understanding to the four-card problem in its original form. Even correct solution of a second problem does not give adequate proof of understanding. The paradoxical choices of some subjects in Experiment II point to the wide prevalence of premature algorithms. When the transfer problem closely resembles the original learning situation, one cannot distinguish true understanding from the application of a rule-of-thumb procedure.

It was hypothesized that realistic content and the "reduced" mode of presentation would facilitate the four-card problem. These predictions were upheld. It was also hypothesized that experience of an easier problem would tend to lead to complete understanding and successful transfer. This prediction was not upheld. But from a Piagetian standpoint, this hypothesis is a necessary corollary to the other two. If it is true that, once he is alerted to the logical structure of the problem by the reduced presentation, a subject will reason in accordance with it, the same subject should not immediately revert to an intuitive pattern of thinking when an analogous, but slightly more difficult, problem is presented.

The theory of formal reasoning may be put in a strong form: "Whenever a problem admits of solution by the application of the propositional calculus, the formal reasoner will attempt to do so." The present findings show that, stated in this form, the theory is false.

One way of reconciling these findings with certain others (Inhelder and Piaget, 1958; Mealings, 1963; Lunzer, 1965, 1968; Roberge and Paulus, 1971; Matalon, 1962, would be to state the theory in a weaker form:

(1) Subjects (other than children) who have attained the level of formal reasoning can be led to apply the propositional calculus correctly.

(2) There exist some situations in which subjects in the former category will spontaneously apply the propositional calculus.

Investigations are currently under way to test the theory in this weaker form and to render its formulation more precise.

References

BRUNER, J. S., GOODNOW, J. and AUSTIN, G. A. (1956). *A Study of Thinking*. New York: Wiley.

INHELDER, B. and PIAGET, J. (1958). *The Growth of Logical Thinking*. London: Routledge.

JOHNSON-LAIRD, P. N. and WASON, P. C. (1970a). A theoretical analysis of insight into a reasoning task. *Cognitive Psychology*, **1**, 134–48.

JOHNSON-LAIRD, P. N. and WASON, P.C. (1970*b*). Insight into a logical relation. *Quarterly Journal of Experimental Psychology*, **22**, 49–61.

LUNZER, E. A. (1965). Co-ordinations et conservations dans le domaine de la géométrie. In BANG, V. and LUNZER, E. Conservations spatiales. *Études d'épistemologie génétique*, **19**, 59–148.

LUNZER, E. A. (1968). Formal reasoning. In LUNZER, E. A. and NORRIS, J. F. (Eds), *Development in Human Learning*. London: Staples.

MATALON, B. (1962). Étude génétique de l'implication. In BETH, E. W., *et al.* Implication, formalisation et logique naturelle. *Études d'épistemologic génétique,* **16**, 69–93.

MEALINGS, R. J. (1963). Problem solving and science teaching. *Education Research*, **15**, 194–207.

PEEL, E. A. (1971). *The Nature of Adolescent Judgement*. London: Staples.

PIAGET, J. (1950). *The Psychology of Intelligence*. London: Routledge & Kegan Paul.

ROBERGE, J. I. and PAULUS, D. H. (1971). Developmental patterns for children's class and conditioned reasoning abilities. *Developmental Psychology*, **4**, 191–200.

WASON, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology*, **12**, 129–40.

WASON, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, **20**, 273–81.

WASON, P. C. (1969*a*). Regression in reasoning. *British Journal of Psychology*, **60**, 471–80.

WASON, P. C. (1969*a*). Structural simplicity and psychological complexity: some thoughts on a novel problem. *Bulletin British Psychological Society*, **22**, 201–84.

WASON, P. C. and JOHNSON-LAIRD, P. N. (1970). A conflict between selecting and evaluating information in an inferential task. *British Journal of Psychology*, **61**, 509–15.

WASON, P. C. and SHAPIRO, D. (1971). Natural and contrived experience in a reasoning problem. *Quarterly Journal of Experimental Psychology*, **23**, 63–71.

RESPONSE ORGANIZATION IN ATTENTION CONTROL AND A PERCEPTUAL SERIAL POSITION EFFECT

GEOFFREY UNDERWOOD

Department of Psychology, University of Sheffield

To determine the influence of response organization factors in selective attention, a comparison was made between the shadowing and monitoring techniques of attention control and a one-trial serial recall technique in which subjects were instructed to remember one message (attended channel) of a dichotic presentation. Detections of semantic targets in the attended and unattended messages from the remembering condition were quantitatively similar to those from the monitoring condition. This suggests that the low detection rates in the unattended message when subjects are shadowing are a function of the higher processing demands of overt response organization required by this task. A serial position effect was also in evidence: the detection probability was enhanced if the target was positioned towards either end of the serial presentation of 16 items. The primacy observed here, common to all three attention control conditions, indicates more efficient perceptual processing and subsequent categorizing of end items than of central, embedded items. The hypothesis is offered that the principles governing the present primacy effects may also underlie primacy in serial position curves of short-term memory studies.

Introduction

Selective attention experiments require that a task be employed which both directs the subjects' attention and allows measurement of the availability of the competing messages. The effectiveness of attention control tasks may, however, be related to the amount of capacity required for their performance. If a fixed amount of capacity is available then the effectiveness of these tasks may be inversely related to the capacity residual from the analysis of the attended message (AM); that is, the capacity free for analysis of the unattended message (UM). Kahneman (1970) has argued that the absence of a continuous response implies a loss of attentional selectivity, but Carey (1971) and Underwood and Moray (1971) contend that the processing demands of the traditional shadowing activity influence the degree of analysis of the UM. The present paper investigates the diminution of the processing capacity allocated to the perceptual identification task consequent upon allocation of capacity to a response connected with the control of attention.

Shaffer (1971) has discussed the relevance of response skills in selective perception, and a number of methods have been employed which have varied the response organization requirements in such situations. Moray and Jordan (1966) used separate response keys for each hand in a dichotic memory span task, and Murray and Hitchcock (1969) used a continuous written response. Both of these techniques provide a complete record of the direction of attention at message selection and eliminate the masking effects of the shadower's own voice, but the processing

demands of overt response organization and response monitoring may also interfere with perceptual identification. Shaffer (1971) has demonstrated that the functionally distinct processes of stimulus perception and response organization can simultaneously occur in a transcription task. However, a situation demanding such parallel processing is not necessarily free from mutual interference effects when a complex perceptual analysis is required. Johnston, Greenberg, Fisher and Martin (1970) have shown that perceptual-motor tracking performances are directly related to the difficulty of a simultaneous perceptual encoding task. Functional independence of simultaneous tasks can only occur when the total fixed capacity (Moray, 1967) is not exceeded.

The issue of shared capacity is highly relevant to the Deutsch and Deutsch (1963, 1967) theory of attention. These authors proposed that all stimuli are initially analysed, and that selection for further processing is accompanied by the inhibition of the processing of non-selected messages. The findings that unattended items are retrievable from buffer storage (Norman, 1969a; Glucksberg and Cowen, 1970) and that items in the UM may semantically interfere with responses to items in the AM in a shadowing task (Lewis, 1970) lend support to this selective response hypothesis. Craik and Levy (1970) have found that semantic relatedness can lead to a recall decrement in a short-term memory (STM) study, perhaps indicating that the interference effect observed by Lewis may not involve processes beyond primary storage after all. To avoid such difficulties of interpretation tasks should be employed which diminish response competition between outputs from the initial analysis system. Deutsch and Deutsch (1967) have criticized experiments in which inhibition of the processing of the UM is encouraged by the shadowing response. Any selective response during a dichotic presentation is vulnerable to this objection, of course, but if an index of the direction of attention is demanded then we can only hope to reduce response interference, and not eliminate it. It is of interest that in situations where dichotic inputs had equal weighting (Moray and O'Brien, 1967; Shaffer and Hardwick, 1969) perceptual analysis was still less than ideal.

In the experiment to be described a serial recall task was used to maintain selective attention to the AM. This design is based on the assumption that the process of attention is necessary for the perceptual analysis of information for recall. An immediate recall task was used to ensure directed attention by Broadbent and Gregory (1963). Although this task removes the constraints of immediate motor organization and monitoring, and fulfils Kahneman's criterion for selectivity in that it is a continuous activity, it only partially meets our requirements because available capacity is allocated to non-perceptual processes (memory organization). Responses to the UM would also be inhibited by the encoding and storage of the AM, according to the Deutsch and Deutsch theory.

A high degree of selective attention may be dependent upon a continuous response to the selected message, but the execution of this response should, on the above argument, also reduce the amount of capacity residual for the experimental task. The present experiment investigates the disruptive aspects of a continuous response by comparing AM and UM analysis over three modes of subject participation: monitoring (no response required), remembering (covert response), and

shadowing (covert and overt response). The detection of semantically defined targets in all possible serial positions of the presentation array was investigated in case qualitative differences might emerge between the three types of response activity. The necessity of an immediate report of a detection as used by Treisman and Geffen (1967, 1968) and Treisman and Riley (1969) is also examined here, and compared with delayed report.

Method

Stimulus materials

A set of 16-pair dichotic lists of randomly selected letter-names (excluding the poly-syllabic "w"), were prepared and checked using the apparatus and method described by Underwood and Moray (1971). The rate of presentation was 2·5 pairs/s, and no dichotic pair contained identical items. Targets for detection were randomly selected digits (excluding "7"), with one such target always replacing one letter in each dichotic list. Target digits occurred in each position in the lists, and in each of the AM and UM lists an equal number of times. The replacement of letters was pseudo-randomly arranged across channels and across serial positions; a target occurred in each position in each channel a total of four times, so 128 lists were recorded, all by a male speaker.

Two sets of recordings were taken from the original master-recording, to check on the consistency between channels. One set had the words "ready, go" (indicating the AM) recorded prior to each list on Channel 1 of the tape, and the other set had this information prior to the lists on Channel 2. Presentation of the recordings was over stereo ear-phones. A morse-type response key, wired to a bulb within sight of the experimenter but out of sight of the subjects, was secured to the table at which the subjects sat. This key was used for immediately reporting detected digits. Shadowing responses were collected on a second tape-recorder.

Subjects

Forty-eight subjects (27 females and 21 males) were drawn from a volunteer subject pool, composed of undergraduates and research students of the University.

Procedure

Subjects were randomly allocated to one of three groups, for the monitoring, shadowing, and remembering conditions. Within these groups they were further allocated to either the immediate or delayed response condition, and to either of the two subgroups differing in the recording channel attended. Repeated measures were taken from the remaining two factors, channel of target and serial position of target.

In all conditions subjects were instructed to detect targets in the messages in both ears, but that they were not to miss any targets in the AM. Half of each group of subjects received the AM recordings over the left ear-phone, and half over the right ear-phone. Subjects allocated to the shadowing condition were first given monaural and dichotic shadowing practice. Practice trials were also given in the remembering condition: three lists of letters, similar to the experimental lists, were presented. Subjects were encouraged to fill in as many spaces in the prepared answer books as possible, guessing if unsure, but leaving blanks where necessary in order to enter each item in its correct serial position. No experimental practice was given to the monitoring group, but one dichotic list was presented to acquaint these subjects with the type of material which they were to encounter. Subjects were asked to equate the loudnesses of the two messages during this period.

Subjects in the immediate response condition were told to press the response key as soon as they thought that they heard a digit in either ear. For an immediate report to score as a
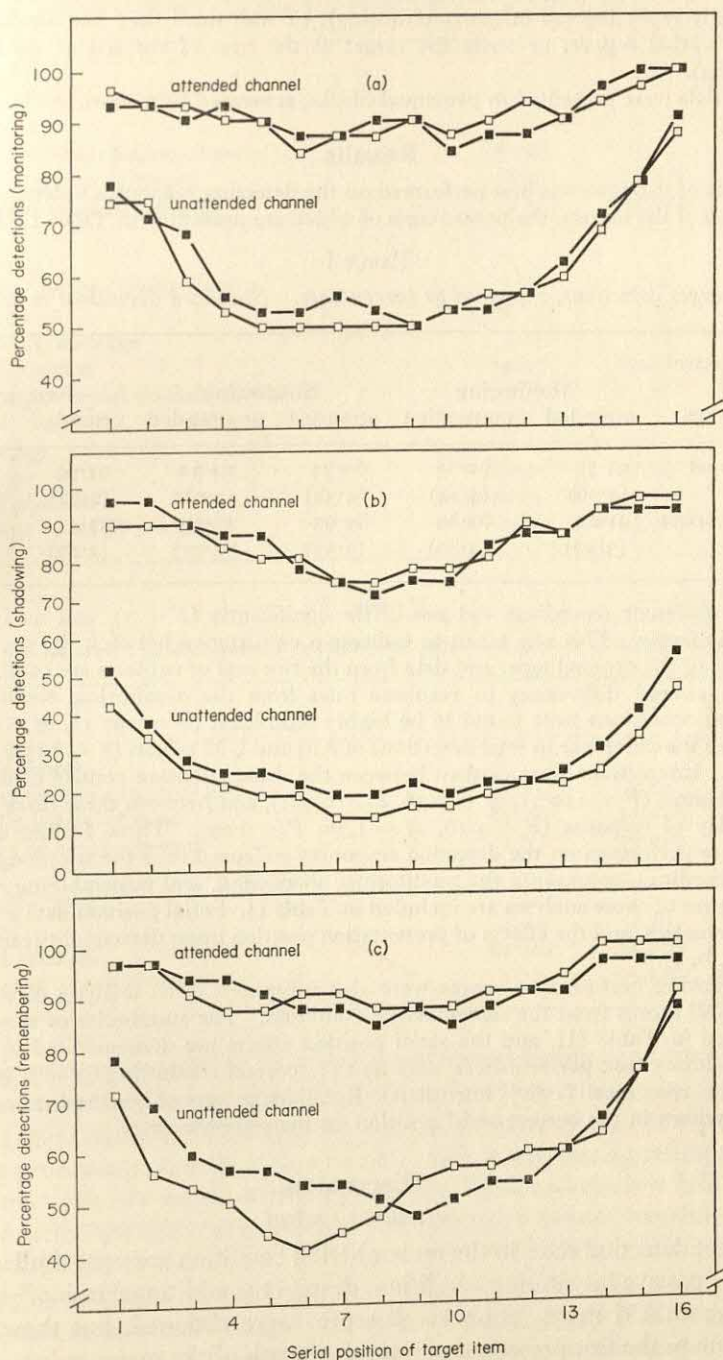
FIGURE 1. Mean percentages of target items detected whilst (a) monitoring, (b) shadowing, and (c) remembering. ■———■ immediate report of target detection; □———□ delayed report.

detection the response had to be executed within the three items following the target. In the delayed response set of conditions subjects either had to: wait until the end of the list before verbally reporting the target (monitoring); or wait until they had shadowed every item before verbal report; or write the target as the first of the list of memory items (remembering).

The 128 trials were presented in two equal blocks, separated by a short break.

Results

An analysis of variance was first performed on the detection responses independent of the serial position of the targets, the percentages of which are presented in Table I. The groups

TABLE I

Mean target detections, expressed as percentages.   Standard deviations in brackets

Attention control response: Target message:	Monitoring		Shadowing		Remembering	
	attended	unattended	attended	unattended	attended	unattended
Delayed report of target:	91·79 (4·26)	60·74 (4·24)	85·74 (4·38)	24·22 (2·07)	91·99 (2·82)	57·62 (3·03)
Immediate report of target:	91·60 (3·37)	62·89 (4·80)	85·93 (4·55)	26·69 (2·07)	91·41 (2·27)	60·16 (6·54)

hearing the different recordings did not differ significantly ($F < 1$), and no interactions involved this factor. This was taken to indicate a consistency between the recordings on each channel of the original tape, and data from the two sets of subjects are combined in the table. The overall differences in response rates from the monitoring, shadowing, and remembering conditions were found to be highly significant ($F = 197\cdot18$, $df = 2,36$, $P < 0\cdot001$), as was the difference in total detections of AM and UM targets ($F = 213\cdot1$, $df = 1,36$, $P < 0\cdot001$). Interactions were evident between the three attention control conditions and the target channel ($F = 110\cdot31$, $df = 2,36$, $P < 0\cdot001$), and between these three conditions and the delay of response ($F = 4\cdot16$, $df = 1,36$, $P < 0\cdot05$). Three further analyses of variance were performed on the detection responses collapsed over the subgroups with the different recordings, comparing the monitoring, shadowing, and remembering conditions. The summaries of these analyses are included in Table II. Serial position data are included in these summaries, and the effects of presentation position upon detectability can be seen in Figure 1 (a, b, c).

The shadowing performance scores were also submitted to an analysis of variance, as were the recall scores from the remembering condition. The summaries of these analyses are presented in Table III, and the serial position effects are demonstrated in Figure 2. Overall subsidiary task performances were 89·13% correct shadowing (3·09% intrusions), and 21·25% correct recall (1·88% intrusions). Recall items were only scored as correct when they were written in the correct serial position on the response sheet.

Discussion

Attention control

The target detection rates in the remembering condition are very similar to those in the more passive monitoring condition, despite the additional task of attempting to learn the 16 AM items. Low recall scores were obtained, but these are presumably due to the fast presentation rate, the length of the memory lists, and high cognitive processing demands of continuous encoding and mnemonic organization whilst detecting targets, rather than to a failure of selective attention. The similarity

TABLE II

Final summary of analyses of variance of detection data

	F	df	P
Monitoring			
Message (attended/unattended)	352·02	1,14	<0·001
Serial position of target	8·77	15,210	<0·001
Message × serial position	3·41	15,210	<0·001
Shadowing			
Report (delayed/immediate)	3·80	1,14	N.S.
Message	2568·33	1,14	<0·001
Report × message	5·15	1,14	<0·05
Serial position	14·57	15,210	<0·001
Message × serial position	2·89	15,210	<0·001
Remembering			
Message	514·50	1,14	<0·001
Serial position	10·51	15,210	<0·001
Message × serial position	2·25	15,210	<0·01

TABLE III

Final summary of analyses of variance on performances in attention control tasks

	F	df	P
Shadowing scores			
Report (delayed/immediate)	1·71	1,14	N.S.
Serial position of item	50·14	15,210	<0·001
Report × serial position	2·69	15,210	<0·01
Serial recall scores			
Report	1·24	1,14	N.S.
Serial position	81·48	15,210	<0·001
Response × serial position	1·36	15,210	N.S.

of detection rates between the monitoring and remembering conditions is possibly due to a neglect of the recall task. However, an experiment by Lindsay and Norman (1969) similarly found that when a discrete perceptual detection task competed for processing capacity with a memory task it was recall performance which diminished. The low recall intrusions from the UM, and the low UM detections compared with AM detections, suggest that subjects were not listening for targets from the UM rather than attempting to remember the letters. If subjects were adopting this strategy and switching their attention to the UM, then it would be expected that most of their targets would be detected in the UM serial positions corresponding to those positions gaining lowest recall. Inspection of Figures 1(c) and 2 indicates that the reverse is the case: the central position items gain both lower recall scores and lower detection rates than do the end position items.

The high detection rates from the AM (Table I), and the significant AM-UM differences (Table II) in all conditions suggest that subjects were focusing their attention on the AM. If subjects had adopted a divided attention strategy then only small differences between the two channels would be expected, following the shared attention studies of Moray and O'Brien (1967) and Shaffer and Hardwick (1969). If we can assume that strategies were consistent between response conditions, then the low detection rates during shadowing must be due to competition between perceptual analysis and overt response organization/monitoring. This conclusion qualifies an earlier criticism (Underwood and Moray, 1971) of the processing demands of the shadowing response.

The bottleneck in multi-channel analysis and subsequent response appears to be partially one of response organization: when the listener is not required to organize and monitor an overt response, then higher detection rates may be observed than at times when he is so required. This is suggested by comparison of the AM-UM differences in detection rates for different response conditions. The decrement produced by shadowing is about 6% in the AM, but is about 35% in the UM (Table I). The results suggest, however, that response requirements are not the sole determinant of the degree of perceptual analysis. All response activities allowed highly significant differences between the AM and UM detection rates. Although items in the UM must gain access to some form of initial storage system, it appears, in the present semantic classification task, that exhaustive semantic analysis of information in the unattended message often does not follow. Lewis (1970) found that verbal reaction times to AM items were subject to interference from semantically similar UM items, and Treisman and Geffen (1967) found a slight semantic context effect in the UM, although a subsequent experiment (Treisman and Geffen, 1968) failed to confirm this effect. The probability of detection of targets in the UM has been found to be inversely related to their semantic complexity (Wilding and Underwood, 1968), perhaps indicating a variable extent of semantic processing of items not receiving attention. Generalizing to the present problem, we might expect that more thorough semantic analysis will be observed with the use of attention control techniques which leave more residual capacity than does shadowing. The present experiment does not appear to be measuring this effect to the same degree however, for only in the end positions of the dichotic lists did detection rates from the UM approach those from the AM. This was true for all three sets of instructions.

Effects of delay

The comparison of the immediate and delayed report conditions suggests that once the listener has categorized an item as a target then it makes little difference whether or not he is required to remember it for a few seconds. This generalization is less valid for the shadowing trials. The necessity to store an item while simultaneously producing continuous overt responses leads to a slight recall decrement, confirming the findings of Kroll, Parks, Parkinson, Bieber and Johnson (1970) and Levy and Jowaisas (1971) that shadowing does interfere with the retention of verbal material.

Interactions between attention control measures and the type of response delay suggest that the structure as well as the quantity of processing demand is important. The key pressing activity of the immediate response disturbs the efficiency of shadowing, and in the remembering task the extra storage load of a delayed-report target interferes with the encoding of late items. In terms of the possible processes involved, it may be that a discrete overt response (key pressing) interferes most with a continuous overt response (shadowing), but that a discrete covert response (committing a target to memory) interferes most with a continuous covert response (learning the list of AM items). Similarly, the accumulative storage of the AM, in the serial recall condition with delayed report of targets, may interfere with the storage and/or retrieval of target digits presented early in the lists.

Serial position effects

Figure 1 shows increased detection rates at the ends of the lists, especially in the UM. In all conditions the serial position of the target was a highly significant variable, and showed considerable interaction with the channel of the target. This is not part of any of the traditional accounts of the serial position effect in STM (Norman, 1969b; Bruce and Papay, 1970) and demands an explanation.

It might be argued that in the shadowing trials primacy results from clear perception of the initial items prior to the onset of the shadowing response. The shadowing lag was of the order of two or three items, thus the items in the first few serial positions would not be masked by the shadower's voice. This explanation does not account for the primacy effects observed during the monitoring and remembering trials.

If the primacy effects in detections were due to increased rehearsal of the initial items then a difference would be expected between the delayed and immediate response conditions. In the delayed condition rehearsal would enhance the primacy; but, with an immediate response, rehearsal of all of the items should be equal and primacy effects minimal. Figure 1 shows that this is not the case; primacy effects are similar in both conditions.

The explanation offered here to account for the detection position effects in detections is based upon the assumption that increased processing time facilitates increased phonetic and semantic decoding; that the longer a stimulus pattern is available for analysis the more probable is a meaningful output from some such word-recognition mechanism as the Logogen System (Morton, 1969). Processing time is not restricted to the duration of sensory experience (Aaronson, 1967) provided that capacity is available. With the rapid presentation of information over two sensory channels residual capacity will be limited, and semantic processing of the secondary channel (UM) may be incomplete. With both the active and postponed perceptual identification strategies described by Aaronson (1968) a "backlog" of items awaiting processing develops. Successive presentations of inefficiently analysed items will be associated with a reduction in the amount of residual capacity available for secondary tasks, and processing of information in the UM will also progressively decline. A perceptual primacy effect operating in this way would also be dependent upon other demands on the available processing capacity. Hence, the effect is most pronounced in the shadowing trials when the subject is fully occupied

by the demands of the tasks. This explanation handles the present data satis-
factorily, although unequal strategical division of attention within the lists (Waugh
and Norman, 1965) could also play a role.

The increase in detections towards the ends of the lists may be due to a reviewing
of the echoic traces of all inputs after the final item has been perceived. At this post-
trial point considerably more capacity would be available for perceptual analysis
than would be during the rapid presentation of information. Additionally, sub-
jective expectancies of targets may influence the detection rate in the later serial
positions. Vigilance studies (Mackworth, 1950) suggest that the increased expectancy
caused by reduced positional uncertainty would increase the probability of detec-
tion of targets presented late in the lists. The slight recency/finality observed
during the shadowing performances (Fig. 2) suggests that either the shadowing
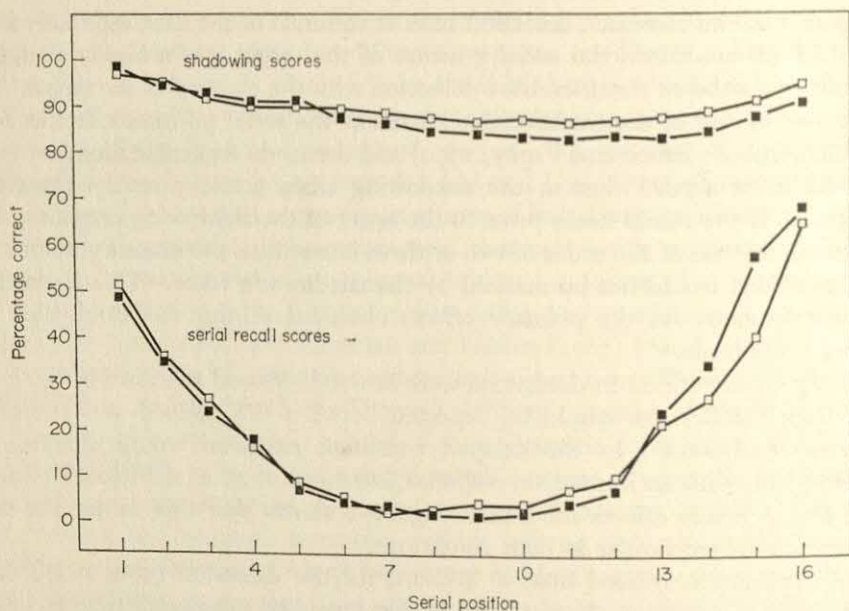
FIGURE 2. Attention control task performances from the shadowing and remembering conditions.
■————■ immediate report of target detection; □————□ delayed report.

voice does not interfere with echoic traces, or that recency effects are not totally due
to the persistence of these traces.

The similarity between the serial position effects in the detection tasks here and
those described by STM studies (e.g. Murdock, 1962) raises the issue of the pro-
cesses common to these tasks, and the relation of these processes to the observed
serial order effects. It has been suggested that the shape of the serial position curve
may be influenced by "perceptual factors" (Aaronson, 1968), and by "variations in
pre-rehearsal storage" (Corballis, 1969), and the present results lend support to this
hypothesis. Interactions between storage and retrieval which are possible causes of
position effects, are eliminated in the immediate response conditions of the
present experiment. Only perception and categorization processes are common
to both the detection and recall paradigms. Massaro (1970) has suggested that

memory strength is directly related to the amount of perceptual processing involved, and it would follow that serial recall effects are also, at least in part, a function of the differential perceptual processing of items. The possibility that common mechanisms underlie perceptual element-position and serial recall curves has also been discussed by Harcum (1967, 1970).

Experiments designed to examine the causes of the position effect have usually controlled only factors involved after acquisition and have, as Aaronson (1968) states, "implicitly assumed that perception of the items presents no problem in sequential recall tasks". The present results and those of Aaronson (1968) indicate the importance of perceptual processes in serial position effects, however, and cannot be explained in terms of rehearsal (Waugh, 1960; Waugh and Norman, 1965; Crowder, 1970), transfer from primary to secondary memory (Waugh and Norman, 1965; Norman and Rumelhart, 1970), proactive interference (Glanzer and Cunitz, 1966; Glanzer and Meinzer, 1967), or retrieval (Shiffrin, 1970). Of course, they do not rule out the possibility of these, and other, processes affecting the shape of the serial position curve in other situations.

In conclusion, it is argued that the effects of the serial position of an item upon its probability of detection as a semantic target and upon its probability of recall in an STM task are both a function of perceptual factors which are common to each paradigm. The primacy effects here, observed without the involvement of memory storage, are based upon a disparity of acquisition potential of individual items, and are interpreted to indicate that these position effects are caused by differential perceptual analysis and encoding.

References

AARONSON, D. (1967). Temporal factors in perception and short-term memory. *Psychological Bulletin, 67*, 130–44.

AARONSON, D. (1968). Temporal course of perception in an immediate recall task. *Journal of Experimental Psychology, 76*, 129–40.

BROADBENT, D. E. and GREGORY, M. (1963). Division of attention and the decision theory of signal detection. *Proceedings of the Royal Society B, 158*, 222–31.

BRUCE, D. and PAPAY, J. P. (1970). Primacy effect in single-trial free recall. *Journal of Verbal Learning and Verbal Behavior, 9*, 473–86.

CAREY, P. W. (1971). Verbal retention after shadowing and after listening. *Perception and Psychophysics, 9*, 79–83.

CORBALLIS, M. C. (1969). Patterns of rehearsal in immediate memory. *British Journal of Psychology, 60*, 41–9.

CRAIK, F. I. M. and LEVY, B. A. (1970). Semantic and acoustic information in primary memory. *Journal of Experimental Psychology, 86*, 77–82.

CROWDER, R. G. (1970). The role of one's own voice in immediate memory. *Cognitive Psychology, 1*, 157–78.

DEUTSCH, J. A. and DEUTSCH, D. (1963). Attention: some theoretical considerations. *Psychological Review, 70*, 80–90.

DEUTSCH, J. A. and DEUTSCH, D. (1967). Comments on "Selective attention: perception or response?" *Quarterly Journal of Experimental Psychology, 19*, 362–3.

GLANZER, M. and CUNITZ, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior, 5*, 351–60.

GLANZER, M. and MEINZER, A. (1967). The effects of intralist activity on free recall. *Journal of Verbal Learning and Verbal Behavior*, **6**, 928–35.

GLUCKSBERG, S. and COWEN, G. N., Jr. (1970). Memory for nonattended auditory material. *Cognitive Psychology*, **1**, 149–56.

HARCUM, E. R. (1967). Parallel functions of serial learning and tachistoscopic pattern perception. *Psychological Review*, **74**, 51–62.

HARCUM, E. R. (1970). Defining shape for perceptual element-position curves. *Psychological Bulletin*, **74**, 362–72.

JOHNSTON, W. A., GREENBERG, S. N., FISHER, R. P. and MARTIN, D. W. (1970). Divided attention: a vehicle for monitoring memory processes. *Journal of Experimental Psychology*, **83**, 164–71.

KAHNEMAN, D. (1970). Remarks on attention control. In SANDERS, A. F. (Ed.), *Attention and Performance III*. Amsterdam: North-Holland.

KROLL, N. E. A., PARKS, T., PARKINSON, S. R., BIEBER, S. L., and JOHNSON, A. L. (1970). Short-term memory while shadowing: recall of visually and of aurally presented letters. *Journal of Experimental Psychology*, **85**, 220–4.

LEVY, C. M. and JOWAISAS, D. (1971). Short-term memory: storage interference or storage decay? *Journal of Experimental Psychology*, **88**, 189–95.

LEWIS, J. L. (1970). Semantic processing of unattended messages using dichotic listening. *Journal of Experimental Psychology*, **85**, 225–8.

LINDSAY, P. H. and NORMAN, D. A. (1969). Short-term retention during a simultaneous detection task. *Perception and Psychophysics*, **5**, 201–5.

MACKWORTH, N. H. (1950). Researches in the measurement of human performance. *MRC Special Report Series*, No. 268. London: H.M.S.O.

MASSARO, D. W. (1970). Perceptual processes and forgetting in memory tasks. *Psychological Review*, **77**, 557–67.

MORAY, N. (1967). Where is capacity limited? A survey and a model. In SANDERS, A. F. (Ed.), *Attention and Performance*. Amsterdam: North-Holland.

MORAY, N. and JORDAN, A. (1966). Practice and compatibility in 2-channel short-term memory. *Psychonomic Science*, **4**, 427–8.

MORAY, N. and O'BRIEN, T. (1967). Signal-detection theory applied to selective listening. *Journal of the Acoustical Society of America*, **42**, 765–72.

MORTON, J. (1969). The interaction of information in word recognition. *Psychological Review*, **76**, 165–78.

MURDOCK, B. B., JR. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, **64**, 482–8.

MURRAY, D. J. and HITCHCOCK, C. H. (1969). Attention and storage in dichotic listening. *Journal of Experimental Psychology*, **81**, 164–9.

NORMAN, D. A. (1969a). Memory while shadowing. *Quarterly Journal of Experimental Psychology*, **21**, 376–86.

NORMAN, D. A. (1969b). *Memory and Attention*. New York: Wiley.

NORMAN, D. A. and RUMELHART, D. E. (1970). A system for perception and memory. In NORMAN, D. A. (Ed.), *Models of Human Memory*. London: Academic Press.

SHAFFER, L. H. (1971). Attention in transcription skill. *Quarterly Journal of Experimental Psychology*, **23**, 107–12.

SHAFFER, L. H. and HARDWICK, J. (1969). Monitoring simultaneous auditory messages. *Perception and Psychophysics*, **6**, 401–4.

SHIFFRIN, R. M. (1970). Memory search. In NORMAN, D. A. (Ed.), *Models of Human Memory*. London: Academic Press.

TREISMAN, A. and GEFFEN, G. (1967). Selective attention: perception or response? *Quarterly Journal of Experimental Psychology*, **19**, 1–17.

TREISMAN, A. and GEFFEN, G. (1968). Selective attention and cerebral dominance in perceiving and responding to speech messages. *Quarterly Journal of Experimental Psychology*, **20**, 139–50.

TREISMAN, A. and RILEY, J. G. A. (1969). Is selective attention selective perception or selective response? A further test. *Journal of Experimental Psychology*, **79**, 27–34.

UNDERWOOD, G. and MORAY, N. (1971). Shadowing and monitoring for selective attention. *Quarterly Journal of Experimental Psychology*, **23**, 284–95.

WAUGH, N. C. (1960). Serial position and the memory span. *American Journal of Psychology*, **73**, 68–79.

WAUGH, N. C. and NORMAN, D. A. (1965). Primary memory. *Psychological Review*, **72**, 89–104.

WILDING, J. M. and UNDERWOOD, G. (1968). Selective attention: the site of the filter in the identification of language. *Psychonomic Science*, **13**, 305–6.

A TEST BETWEEN THE SELECTIVE ATTENTION AND STIMULUS GENERALIZATION INTERPRETATIONS OF THE EASY-TO-HARD EFFECT

J. SWELLER

University of Adelaide

The selective attention and stimulus generalization interpretations of the easy-to-hard effect were tested by training one group of rats on an easy brightness discrimination and "reversing" them on a more difficult brightness discrimination. A control group, initially trained on the difficult discrimination, was reversed at the same time as the experimental group. The experimental group learned the reversal more rapidly and this was interpreted as supporting the selective attention interpretation.

Introduction

If an animal is trained on an initial discrimination and then transferred to a more difficult discrimination on the same dimension, learning this second task is characteristically more rapid than if all training is given on the difficult discrimination only. This is known as the easy-to-hard effect. It was first demonstrated by Gubergritz (in Pavlov, 1927) using classical conditioning of the salivary reflex of a dog. Lawrence (1952) obtained the effect using instrumental conditioning of rats and suggested that selective attention could account for it. If as well as learning the choice response, animals must learn to attend to the relevant dimension, then learning should be more rapid if the conditions are such that the relevant dimension is isolated early in training, as will be the case if an easy discrimination is used.

Lawrence put forward this account on the assumption that excitatory and inhibitory stimulus generalization gradients of the type postulated by Spence (1936) could not account for the effect (see Lawrence, 1955). He pointed out nevertheless, that if the gradients became steeper with increased training (rather than being parabolic in shape as proposed by Spence) then the effect could be explained without recourse to attention. This explanation was accepted by Logan (1966) who showed that a modified Spence model could not only predict the easy-to-hard effect but could also predict the conditions under which it would be larger or smaller. He assumed that bell-shaped excitatory (E) and inhibitory (I) stimulus generalization gradients steepen while building up about the positive and negative stimuli respectively. The net E or I at any point on a stimulus continuum is calculated by obtaining the difference between E and I at that point. If two stimuli are far apart on the continuum (an easy discrimination), then after training on these stimuli, the net E and/or I will be greater at two points closer together on the continuum (a difficult discrimination) than if all training had been carried out using these latter stimuli only. The easy-to-hard effect is hence predicted.

Singer, Zentall and Riley (1969) have provided a direct test between the selective attention and stimulus generalization gradient interpretations. They reasoned that if easy discrimination training using three stimuli such as black, grey and white presented in pairs with the middle value present on all trials is substituted for easy discrimination training using two stimuli (e.g. black-white) then equal amounts of generalized E and I should occur at the two stimuli used for the hard discrimination (e.g. light-grey and dark-grey), eliminating the usual facilitatory transfer effect. In order to avoid differential E and I for the group trained on the hard discrimination only, the comparison was made from the first trial of the difficult discrimination for both groups. This necessitated the inclusion of a third, learning-to-learn control group initially trained on another dimension before switching to the hard discrimination.

Singer et al.'s (1969) results supported the stimulus generalization interpretation in that the easy-to-hard effect was not obtained using a three-stimulus easy discrimination group (i.e. there was no difference between this group and the difficult discrimination only group other than that seemingly caused by a learning-to-learn effect) but was obtained using the conventional two-stimulus easy discrimination groups. The attentional interpretation predicts that three-stimulus easy discrimination training should have the same effect on dimensional salience as two-stimulus training and hence was not supported.

In order to increase the generality of Singer et al.'s conclusions, a further test of the two theories is proposed. Advantages of the proposed test are that it is considerably simpler and also that its essential logic does not involve acceptance of the null hypothesis at any point as is the case with Singer et al.'s experimental design. For the normal two-stimulus paradigm, if net E and I are relatively large at the hard stimuli for the "easy" group, then by making the usual easy-to-hard shift a "reversal" shift as well, and by simultaneously reversing the subjects in the "difficult" group a reverse easy-to-hard effect is predicted: i.e. the easy group should learn more slowly because it has a larger net E and I to reverse. The opposite result would constitute strong evidence against the stimulus generalization account. It might on the other hand, be explained by the attentional interpretation given the assumption that increased attention to the relevant dimension is beneficial irrespective of whether or not the cues are reversed.

Method

Subjects

Twelve experimentally naive male hooded rats, approximately 120 days old at the beginning of the experiment were used.

Apparatus

The apparatus was constructed of flat grey painted wood with clear perspex lids. It consisted of a 14·0 cm long, 6·5 cm wide, and 9·0 cm high starting box separated by a guillotine door from a 26·5 cm long, 20·5 cm wide, and 9·0 cm high runway ending at a 7·5 cm wide airgap. The runway then continued as two parallel alleyways 20·5 cm long and separated by a 1·5 cm wide partition. Two top hinged, removable stimulus doors, 7·5 cm high and 9·0 wide could be placed 12·5 cm from the airgap in each alleyway. The runway continued for 20·5 cm past the two alleyways before opening into the goal-box which was 23·0 cm long, 20·5 cm wide and 15·0 cm high. A 60 W globe was located 10·0 cm in

front of the stimulus doors, immediately above the perspex lid. For training, black, dark grey, light grey and white painted aluminium doors were available. Their brightness readings using a spot photometer in the normal experimental lighting conditions were 0·5, 1·2, 1·7 and 2·0 log foot-lamberts respectively.

Procedure

Pretraining. During the 10 days of pretraining, animals were tamed, reduced to and maintained at 85% of their ad lib weight, and trained to run through black-white vertical striped doors for food. All animals experienced locked doors on some occasions. Manual guidance was given throughout in order to equalize experience with both positions.

Training. Animals were given 10 non-correction trials a day in groups of three giving an inter-trial interval of approximately 3 min. The position of the stimuli were varied according to a Gellermann series. The negative stimulus door was always locked. Correct responses were rewarded by 30 s access to wet mash. After incorrect responses animals were retained in the apparatus for 10 s. The criterion of learning was 18/20 correct responses over two days.

Design. There were two groups of 6 animals each. Those in the easy group were trained on the black-white discrimination until they had reached criterion. They were then reversed to criterion on the light grey-dark grey discrimination: e.g. if white had been positive, after transfer dark grey became positive. Half of the animals initially had white positive, and the other half had black. The animals in the difficult group were given initial training using the light grey-dark grey stimuli. Each animal was randomly paired with an animal in the easy group and was reversed on the same day as its partner was shifted. Training continued to criterion. Half of the animals were initially trained with light grey positive and the other half with dark grey.

Results

Mean results are given in Table I. Days to criterion include the two criterion days. Statistical tests yielded the following results:

(1) the easy group had significantly fewer correct trials by a *t*-test on the first shift day than the difficult group, $t = 5·43$, $df = 10$, $P < 0·001$;

(2) the easy group took significantly fewer days by a *t*-test to reach the shift criterion than the difficult group, $t = 4·31$, $df = 10$, $P < 0·005$.

TABLE I

	Days on initial task	Correct trials on first shift day	Days to shift criterion
Easy group	6·33	1·67	16·17
Difficult group	6·33	4·33	22·00

Discussion

The results indicate that the performance of the easy group was significantly worse than that of the difficult group on the first shift day but despite this the easy group took significantly fewer days to reach criterion. Since according to the stimulus generalization interpretation, the performance of the easy group should have been inferior at all stages of learning, the former result is predicted but the

latter is not. The most obvious explanation of both results is that the shift learning of the easy group was mediated by a higher level of attention to the correct dimension resulting in an initial greater difference in the response strengths to the two cues and subsequently followed by a more rapid rate of change.

That the conditions were similar to those normally pertaining in successful easy-to-hard effect paradigm experiments, can be seen from the first shift day results. Had there not been a reversal shift, an identical result would have indicated that the easy group had performed significantly better rather than significantly worse than the difficult group, giving the normal easy-to-hard effect.

The present results are probably analogous to those obtained by Mackintosh and Little (1970) using pigeons. They found that subjects "reversed" to a difficult task from an easy task (as was the easy group of the present experiment) eventually performed more accurately than control subjects given all their training on the difficult task with no reversal.

The contradiction between the theoretical position supported by the results of Singer *et al.* (1969) and the present results requires an explanation. Three-stimulus learning may have forced animals to respond to the absolute properties of the stimuli rather than to relationships. (Singer *et al.* provide evidence that this may have occurred.) This mode of responding may generalize to all stimuli on the same dimension. If one assumes that abolute responding is more difficult for some or all animals than either relational responding or a combination of relational responding and absolute responding (as will be the case if some animals must adopt a mode of responding not naturally adopted), then this may negate any advantage accruing to the easy-to-hard three-stimulus group due to a higher level of attention to the relevant dimension.

References

LAWRENCE, D. H. (1952). The transfer of a discrimination along a continuum. *Journal of Comparative and Physiological Psychology*, **45**, 511–6.

LAWRENCE, D. H. (1955). The applicability of generalization gradients to the transfer of a discrimination. *Journal of General Psychology*, **52**, 37–48.

LOGAN, F. A. (1966). Transfer of discrimination. *Journal of Experimental Psychology*, **71**, 616–8.

MACKINTOSH, N. J. and LITTLE, L. (1970). An analysis of transfer along a continuum. *Canadian Journal of Psychology*, **24**, 362–9.

PAVLOV, I. P. (1927). *Conditioned Reflexes* (translated by G. V. Anrep). London: Oxford University Press.

SINGER, B., ZENTALL, T. and RILEY, D. A. (1969). Stimulus generalization and the easy-to-hard effect. *Journal of Comparative and Physiological Psychology*, **69**, 528–35.

SPENCE, K. W. (1936). The nature of discrimination learning in animals. *Psychological Review*, **43**, 427–9.

SYMMETRY, GESTALT AND INFORMATION THEORY: A CRITIQUE

ALAN M. WING

*Psychology Department, McMaster University,
Hamilton, Ontario*

A criticism of a recent analysis of matrix pattern perception is advanced. It is pointed out that a fair test of the equivalence for performance of different forms of redundancy must be framed in terms of the measure of transmitted information.

Deregowski (1971) reports an experiment in which spatial grids, with a cross or blank in each cell, were used to define visual patterns. These differed on a number of factors but not on the information measure derived from filled and empty cell probabilities. Constraints were imposed on the patterns, as will be described below, which introduced redundancy. That is, the uncertainty associated with any pattern, or matrix, was less than if all cell entries had been defined on an equiprobable basis.

On each trial, after the subject had viewed a tachistoscopically exposed matrix for 2 s, he was required to reproduce it on a sheet of paper. The general theoretical interest was in making a comparison of predictions of the traditional information approach with the Gestalt law of figural goodness and symmetry. In this article it will be shown that the response measure used by Deregowski (the number of wholly correct responses) was inappropriate on the basis of certain information theory considerations.

The method of stimulus construction used gave rise to four sets of matrices that were equally predictable in the sense that, the amount of uncertainty or information in each of the sets was equal, but the forms which the redundancy took were different across sets. The types of redundancy in the 4 × 4 matrices were defined by operations on a 4 × 2 submatrix, with 4 filled and 4 empty cells, to produce a second 4 × 2 submatrix placed to the right of the first. The resultant 4 × 4 matrix was then one of the following types, in which (1) the second 4 × 2 submatrix was the mirror image of the first, (symmetry about the vertical axis), (2) the matrix in (1) was rotated 90 degrees clockwise, (symmetry about the horizontal axis), (3) the second 4 × 2 submatrix was a repetition of the first, (repetition about the vertical axis), (4) the matrix in (2) was rotated 90 degrees clockwise, (repetition about the horizontal axis). Further manipulations involved emphasis of the symmetry/repetition axes, either, by filling in the second submatrix with crosses of a different colour from those in the first, or, by darkening the line of the grid forming the horizontal/vertical axis.

The results obtained, based on the numbers of completely correctly reproduced matrices, were taken as showing an effect of symmetry vs. repetition forms of

redundancy. Deregowski further claimed that rotation of axis had a significant effect on performance, as did the introduction of two-colour patterns.

Garner (1962) has given a very complete and thorough discussion of the usefulness and the limitations of the information measure in specifying the stimulus. Very briefly, the point of the information measure, in the present context, is that it can be specified on the two aspects of the stimulus—number of stimulus elements and their probabilities of occurrence—which define the unpredictability of the matrix in advance of observation. Discrimination performance may then be determined relative to that level of predictability.

The uncertainty associated with the stimulus matrix is given by,

$$H_S = \underset{ij}{\Sigma\Sigma} - \pi_{ij} \log_2 \pi_{ij} \tag{1}$$

where i indexes the 16 cells in the matrix and j, the stimulus elements, (cross or blank), in each cell. The probability associated with stimulus, j, in the ith cell is given by π_{ij}, so that $\underset{j}{\Sigma}\pi_{ij} = 1$.

For each cell, i, the subject is required to make the response k. If this response has associated marginal probability, p_{ik}, the response uncertainty is given by,

$$H_R = \underset{ik}{\Sigma\Sigma} -p_{ik}\log_2 p_{ik} \tag{2}$$

and $\underset{k}{\Sigma}p_{ik} = 1$.

Thus, considering the system of stimulus and response, a maximum reduction in uncertainty that could result from one stimulus-response pair for any given set of marginal stimulus probabilities, π_{ij}, and marginal response probabilities, p_{ik}, is defined by,

$$H_{\max} = H_S + H_R. \tag{3}$$

In practice, the maximum uncertainty reduction may not be attained and the actual amount of information transmitted will be given by the expression,

$$I = H_{\max} - H_{\text{obs}}. \tag{4}$$

The second term, H_{obs}, represents the joint uncertainty of the data and is given by,

$$H_{\text{obs}} = \underset{ijk}{\Sigma\Sigma\Sigma} -\pi_{ij}p_{ijk}\log_2\pi_{ij}p_{ijk} \tag{5}$$

where p_{ijk} is the probability of response, k, conditional on stimulus, j, in the cell, i.

Thus, in assessing the effect of various forms of redundancy in the stimulus, consideration of the transmission of information through the subject requires analysis of the whole of the stimulus-response data matrix. Analyses based on wholly correct responses only are wasteful of data. But, of greater concern, without strong assumptions with respect to response bias and the form of the joint data matrix, one cannot interpret such as evidence, for or against, the effects of different forms of redundancy on performance.

Suppose a more complete analysis were to reveal that different types of redundancy do result in equivalent information transmission. The major theoretical

question would then be: how are different forms of redundancy assimilated by the subject with equal effectiveness? In either event, the argument that has been presented here is *not* that information theory might account for data obtained from such an experiment, but rather that the phenomenon to be explained might be an absence of effect of the type of constraint used in the reduction of uncertainty.

References

DEREGOWSKI, J. B. (1971).   Symmetry, gestalt and information theory.   *Quarterly Journal of Experimental Psychology*, **23**, 381–5.
GARNER, W. R. (1962).   *Uncertainty and Structure as Psychological Concepts*.   New York: Wiley.

A REJOINDER TO "SYMMETRY, GESTALT AND INFORMATION THEORY: A CRITIQUE"

JAN B. DEREGOWSKI

Department of Psychology, University of Aberdeen

That the data obtained in the experiment in question (Deregowski, 1971) could have been further analysed using information theory techniques—I admit. That they should have been so analysed—I deny. I do so for the following reason: The experiment was intended to contrast two approaches to the problem of symmetry, therefore the critical measure used needs to be acceptable to both approaches; a perfectly correct response fulfils this desideratum.

When two independent theories of description of a characteristic, each with its own scaling system, are being compared and one does not know how the scales relate to each other, along their entire lengths, nor how each of them relates to the value of the absolute characteristic (if one knew this one could, obviously, deduce the relationship between the scales) the only legitimate comparisons possible are at those values of the absolute characteristic at which both of the relationships between the scales are known. At any other value the unknown relationship between the scales makes such a comparison impossible. (Use of *both* scales may enable one to discover the relationship between them—this, however, was not the issue.)

One can view this argument from another stance. If the assessment of correctness of a reproduction of a pattern be operationally defined by a measurement theory, and two such theories exist and the relationship between them is not known, then such theories cannot be compared, by deriving a set of measurement of stimuli *in terms of one of these theories.*

In the experiment discussed, both the Gestalt theory and the information measurement concepts are considered. One does not know how the two theories are related but one knows that a perfect reproduction is regarded as a correct response by both the Gestalt and the Information theory standards.

Once one departs from this point and considers other than perfect responses the two measures differ. Consider Figure 1 showing possible results.
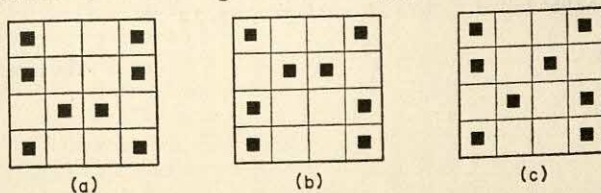
FIGURE I. (a) Stimulus, (b) and (c) possible responses. Four points are misplaced in response (b) only two in response (c), yet response (b) retains its symmetry about the vertical axis (an attribute of Gestalt "goodness") while response (c) loses this. In fact, response (b) is the inversion of the stimulus and response (c) is skew-symmetrical and hence still has a degree of redundancy.

The symmetrical displacement of two pairs of dots [Fig. 1(b)] would be hailed by a Gestaltist as evidence for preservation of figural "goodness", as a superior response than that shown in Figure 1(c), where only two dots have been misplaced. An information theory analyst would, on the other hand, consider Figure 1(c) to be less erroneous than Figure 1(b). The measures thus diverge and neither the use of both of them nor abandonment of one of them and adoption of the other (a course implied in Wing's critique) offers a solution to the problem.

Hence it appears to the writer that the analysis used by him cannot be said to be wasteful of data since it is the only way of analysing data available at the moment, nor can it be said that the conclusions derived from this analysis are invalid.

This is the main reason for the submission that Wing's critique is confuted.

Two further points ought perhaps to be considered. (1) The stimuli were constructed by using random number tables and allocating an "X" to a cell whenever an odd number was encountered. This means that not all the stimuli had four filled and four empty cells in the familial submatrix. [The actual frequencies of filled cells were: 2, 3, 4 (twice), and 5 (twice).] Since each subject responded to five stimuli each of which was from a different familial group, it is unlikely that any response bias involving the number of filled or empty cells was present. (2) The test series used were short consisting of five stimuli. Hence it seems also unlikely that the subjects became aware, in the course of the experiment, of the probabilities involved in construction of the stimuli. This is important as shown in Green and Courtis (1966) and by Garner (1962, p. 186), in his discussion of redundancy which "increases discriminability of the stimuli actually used, but this increased discriminability can be of value only if it is perceived".

In the case in question the comparison of other stimuli with the random stimulus leaves little doubt that redundancy due to symmetry and repetition was made use of, whereas there is little to suggest that redundancy associated with the frequency with which the filled cells occurred was in any way involved.

I am indebted to Mr K. J. Gilhooly for acting as an *advocatus diaboli*.

References

DEREGOWSKI, J. B. (1971). Symmetry, gestalt and information theory. *Quarterly Journal of Psychology*, **23**, 381–5.

GARNER, W. R. (1962). *Uncertainty and Structure as Psychological Concepts*. New York: Wiley.

GREEN, R. T. and COURTIS, M. C. (1966). Information theory and figure perception: The metaphor that failed. *Acta Psychologica*, **25**, 12–36.

THE ADDING OF VISUAL INFORMATION TO PREVIOUSLY RECORDED VIDEO-TAPES: A FUNCTIONING SYSTEM

J. D. ARMSTRONG AND R. C. WORSEY

Blind Mobility Research Unit,
Department of Psychology, University of Nottingham

Introduction

The Sony Corporation now market a number of pieces of equipment which allow considerable versatility in both the recording and the presentation of video material.

One useful facility is provided by the Sony Video Camera Wiper equipment which allows a split-screen presentation of video information from two separate video cameras. For example, one camera could view a series of events whilst the other might be focussed on a digital clock. The Video Camera Wiper equipment allows a composite picture to be produced which can then be recorded for future reference.

Work carried out in this Department involves the video recording of blind people in street situations. Recordings, which are made on a portable video-recorder, are brought back to the laboratory for replaying on a standard VTR. Normally, it is not possible to add any insert (e.g. a timing strip) to a previously recorded tape since the Sony Video Camera Wiper will only accept two real-time inputs for simultaneous split-screen presentation.

A system, which has been developed to allow the mixing, and subsequent recording of previously recorded material with a real-time video signal, is described here.

The Equipment

The mixing of two video signals is only possible when they are both controlled by the same master synchronizing source. In the case of the real-time mixing of the two camera inputs, both cameras derive their synchronization from the same sync pulse generator in the wiper unit.

However, when a recorded tape is the source of one of the video signals to be mixed, the sync pulse relating to that signal is already present on the tape. If this recorded material is to be mixed with a real-time video signal, then the sync pulses already on the tape must be used as a master synchronization source rather than the sync pulse derived from the wiper.

A system has been developed to extract and suitably modify the synchronization pulses from the tape so as to provide master synchronization for both the wiper and the camera which provides the wipe insert picture.

The standard system used by the Blind Mobility Research Unit consists of a Sony Video-recorder (CV 2100 ACE) which plays back tapes previously recorded on a Sony Portable Video-recorder (DV 2400 CE) and the video picture is viewed on a Sony Monitor CVM 90 UB). Although the Field Pulse is deriveable directly from the video-recorder, it is not really suitable for slaving purposes. Further, no Line Pulse is extractable from the machine. However, both Line and Field Pulses are available after the sync-separator in a normal monitor (including the CVM 90 UB) and these may be extracted via a capacitor and then modified with a Schmitt trigger and clipper circuit to be used as an external sync-source (see Fig. 1).
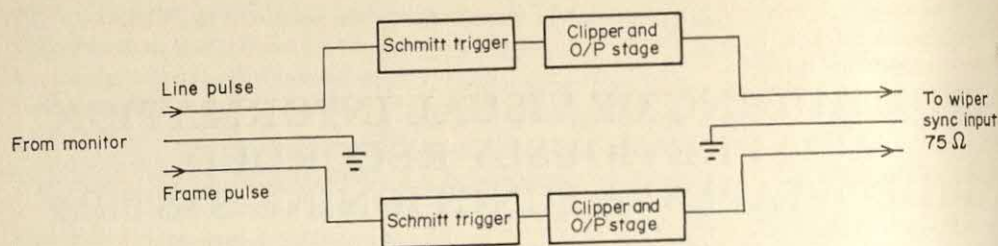
FIGURE 1

The extra equipment needed for mixing the video signal from the recorded tape with the real-time signal from a camera consists of a Sony Video Camera Wiper (CMW 100 CE) and a Sony Video Camera (AVC 3200 CE). The inter-connections of the equipment are shown in Figure 2.
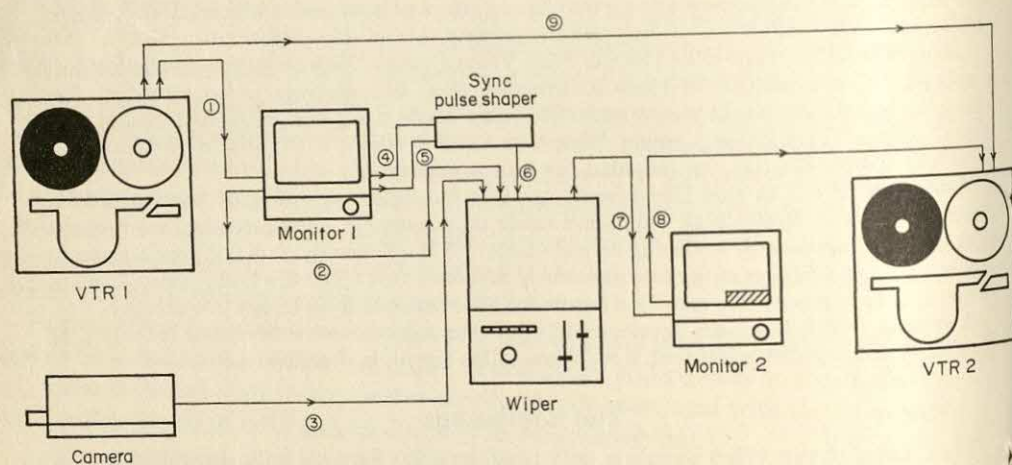
FIGURE 2

The previously recorded video tape is played back on VTR 1 and displayed, via connection 1, on Monitor 1. This video signal is mixed by Sony Video Camera Wiper with the signal derived from the video camera (connections 2 and 3).

The Line Sync and Field Sync are derived from the Monitor 1 (connections 4 and 5) and, after shaping, from the external sync source for the wiper (connection 6).

The composite picture is then displayed, via connection 7, on Monitor 2 before being recorded on VTR 2 (connection 8). The sound track is recorded via the direct connection 9 from VTR 1 to VTR 2. As with any re-recording process, some loss in video quality must be expected. However, provided that the playback and recording heads of VTR 1 and 2 and the recording tapes are in order then little deterioration is obvious during one process. Special attention must also be given to ensure that the video and audio signal levels are correct at all stages of the process.

Applications

The information which could usefully be added to the previously recorded tape is obviously a function of experimental demand. Two types of insert have been made in our current research programme.

The first insert, and probably the most universally useful one, is a digital clock. The progress of a blind pedestrian along a test route is recorded on video tape and on replay the read-out of a two register digital clock is added at the bottom of the main picture.

The second type of insert is of physical data referring to the behaviour being observed.

Information from an accelerometer attached to a pedestrian's body is telemetered back to the portable video-recorder and recorded on the sound track of the tape as an audio signal. Later, in the laboratory, the audio signal is transferred to a high-speed pen recorder which is viewed by the camera in the split-screen equipment. In this way, the video picture of the pedestrian is played back simultaneously with an insert of a visual representation of his body acceleration and the composite picture is then recorded for analysis at a later time.

The actual electronic circuitry of the sync pulse shaper is not given here but is available, on request, to any bona fide University Department.

BOOK REVIEWS

SANDERS, A. F. (Ed.). *Attention and Performance III.* Amsterdam: North Holland Publishing Company. 1970. Pp. viii + 442. (Price not known.)

The published reports of symposia are notoriously unsatisfactory. The papers are frequently of only temporary interest, looking extremely frail and wilting in the rigid embrace of hard covers: and because the editor can hardly exclude a paper which was given at the actual meeting, it is too often true that the contents fall below the standard required by a journal. There are, however, certain exceptions to this depressing rule, and the first two symposia on Attention and Performance provided them. Now the third volume has done it again, and suggests that there is some magic recipe known only to Dutch psychologists, which allows them to produce these consistently important volumes in a form that has led so many other people to disaster. This volume is the report of a symposium held in Soesterberg in 1969. Although it contains 30 papers given on that occasion, every single one of them is worth reading, and some are likely to be quoted for years to come.

The papers are divided into five sections, dealing with reaction processes, selective attention, subjective probability and performance, memory mechanisms and information processing, and studies on sustained attention. Each of these sub-divisions has a rather different flavour: the first splits into a number of factual contributions to topics having little to do with each other. One cluster of papers deals with multiple stimuli, one with factors altering the importance of response probability in reaction time, one with the trade-off between speed and accuracy, and a last small cluster with search processes in vision and memory. In the field of selective attention, on the other hand, there is a constant feeling that each paper should be compared and contrasted with the others, since they all deal with related topics: and variety is rather given by a range of approach between broadly sweeping theoretical accounts of attention on the one hand, new data on the other, and physiological correlates as a sideline. The section on subjective probability is again heterogeneous, perhaps the only point which marks papers in this group off from the rest of this book being the absence of information processing models in them. One paper is a review of probability learning in the multiple alternative case, one a review and fresh results on the extent to which people can produce random sequences of responses, one a choice reaction time experiment fitted to an abstract mathematical model of reaction time, and one a study of the relative effects on step tracking of the number of possible steps and of variation of sequential structure. With the fourth section on memory we are back in the realm of information processing models, with theory and results on echoic and iconic storage, retrieval, and decision processes in memory. Lastly, the fifth section on sustained attention contains papers which all deal with vigilance or with blocking in continuous work, have a high factual content and mostly look to physiological data rather than to flow diagrams within the system.

What then can one say of the book as a whole? Clearly there is no dominant theme, nor does one even feel the convergence of different lines of thought before some new synthesis emerges. Rather, the impression left by the book is of an impressive body of able workers, speaking the same language and sharing certain technical and statistical techniques, but tackling a large array of fairly separate problems. There is no sign of exhaustion or boredom in this field, nor of any drop in the ability of those working in it. Part of the Dutch magic in producing these good symposia is undoubtedly skill in the selection of the speakers: but so high a percentage of successes speaks well of the quality of workers in the field generally. One is left, however, with an impression of a number of quite disparate problems each yielding a little to careful attack, rather than a co-ordinated assault on some key question of science. In keeping with this general style, perhaps the very best and most striking papers are those which are most factual and most selective in their coverage. One can instance

Rabbitt's division of perceptual failures from errors of response selection and execution, by looking at speed and accuracy trade-offs in different situations: or Marcel's demonstration that the now classic results of Sternberg on search through memory apply only for stimuli characterized by particular combinations of features. One thinks also of Anne Treisman's series of experiments on dichotic listening, in which success or failure, in separating the two ears perceptually, varies with a large number of surprising variables: or von Wright's comparison of selection in the Sperling situation using various different features of the stimulus as a basis for making the selective response. The broader reviews and general theoretical statements are useful surveys, but much less incisive and exciting. Perhaps this is as well: vast conceptual syntheses may be emotionally satisfying to the believer, but have a way of getting their effect by ignoring the awkward and curious. Thus it is as well to be reminded by Sanders himself of the role of motor preparation in response time, or by Trumbo of the fact that we still do not understand serial tracking.

This book is therefore a statement of work in progress: and not calculated to attract those who want to find out why information processing is important, or what the general ideas are in the field. But for those who already know why they want to study this topic, it is essential reading.

D. E. BROADBENT

CORCORAN, D. W. J. *Pattern Recognition.* Harmondsworth: Penguin Books Ltd. 1971. Pp. 223. £0·50.

This is a good tempered and clearly written book, and is considerably more original than most elementary texts. It has some very sensible things to say about work on absolute judgements and deftly disposes of the information theory approach to visual form perception. It also treats illuminatingly serial versus parallel processing, though limitations of space prevent this topic being developed in sufficient depth. Early theories of form perception in animals are reviewed but the weaknesses in these "encoding" theories are not sufficiently brought out and their purveyors (including the reviewer) deserve more castigation than they receive.

Corcoran argues that the processes involved in pattern recognition are passive rather than active, and a considerable portion of the book is devoted to experiments of his own and others showing that in certain simple cases involving a limited number of dimensions of stimulus variability, the outcome of judgements in noise can be fitted by a model in which decisions are taken about the presence versus absence of particular features in the stimulus and the final judgement is arrived at by combining independently the results of these separate decision mechanisms. Unfortunately, the situations in which this has been found to be true are so simple that it can be doubted whether the results throw much light on what characteristically occurs in perception of the real world. Corcoran starts from the assumption that all that is involved in pattern recognition is some sort of classification process, but this view blinds one to the richness of what actually occurs in perception. If perception were merely classification, then it becomes difficult to give any account of perceptual learning or of the processes that enable us to give an accurate description of a completely new object presented for the first time. The problem of how it is that a two-dimensional input to the retina is mapped onto a description that explicitly represents three-dimensional structures and relationships is completely ignored, and no hint is given that perception may involve complex inferential processes and that the process of analysis may be guided by existing knowledge.

The view of perception as classification seems to have influenced not only Corcoran's own thinking but also the references he cites. For example, his chapter on pattern recognition by computer refers only to work based on the idea of classification: there is no description of Roberts' early work on the analysis of three-dimensional scenes nor of the more recent work of Guzman and Minsky; work involving picture grammars undertaken by Clowes, Ledley and Narasimhan also receives no mention. Similarly, there is no reference to such psychologists as Bartlett, J. J. Gibson and Helmholtz: although the Gestaltists are mentioned in passing, there is no discussion of many of the phenomena to which they called attention such as Wertheimer's laws of grouping.

No book of this size could, of course, adequately deal with all the problems of pattern perception particularly as work on speech recognition as well as work on vision is included. However, it could be argued that the issues with which the book deals are the tangential ones rather than the deep ones and that to this extent it may be seriously misleading for students.

Within its limitations, however, the book does a very good job. Experiments are clearly reported and the theoretical issues with which it deals are well expounded. Provided it is borne in mind that pattern recognition involves many problems that are not touched upon by Corcoran and that these may in the long run be the more rewarding problems, students and more senior workers should profit from reading this book.

N. S. SUTHERLAND

SCHRIER, A. M. and STOLLNITZ, F. (Eds.) *Behaviour of Nonhuman Primates Vol. III.* New York: Academic Press Inc. 1971. £4·45·

The third volume of this series establishes it more firmly as an essential source of reference for a wide variety of workers who use primates in behavioural research. As in earlier volumes, the reviews which it contains are not all clearly related, except in so far as all are concerned with some aspect of primate behaviour. R. A. Hinde provides a particularly thorough review of work on the development of social behaviour up to 1969. The value of this chapter is obvious from the initial tables, which summarize what is known of the order and ages at which different stages of behavioural development occur, and at the same time point out very clearly the inaccuracy of much of what has been recorded, and also the gaps which remain to be filled. Even in the case of the rhesus, we still cannot make the sorts of comparison with human development which would be really informative. However, the number of quantitative studies is growing at an impressive rate, and, what is more important, methods of analysis and experimental design are becoming more sophisticated and revealing. Hinde's own indices of the relative importance of mother and infant in the maintenance of contact are excellent examples. A number of important problems of human development can now be seen to be open to experimental analysis in primates. Analysis of early sex differences in infant social behaviour promises to illuminate the genesis of similar differences in man, whilst the discovery of very persistent behavioural consequences in rhesus of brief separation from the mother may help us not only to understand the effects of human deprivation of affection but eventually to test treatments for it. The subject is also not without quite unexpected findings, of which the use of infants in some primate societies to confer immunity from attack and to gain status, is perhaps the most surprising.

Bernstein's contribution is intended as a demonstration of the validity and usefulness of a new technique for the characterization of the social behaviour of different primate species. It is concerned entirely with his own work. I remain unconvinced that the analysis of behaviour derived from samples of individual behaviour lumped across all members of a group, into such broad categories as travel or agonistic responses, is the best approach. However, it is clear that some concise way of describing the general *facies* of social behaviour in different species is needed, which can be based on quantitative data. Bernstein's further analysis of agonistic encounters into ones which involve either physical contact or more distant threat points the direction of development which the technique will no doubt follow.

The final two articles review the psychophysics of vision and hearing in primates. I found the first (by DeValois and Jacobs) quite outstanding, both in its presentation and breadth of cover. The account of colour vision is particularly helpful in bringing order to a confused literature, and a difficult subject. The distinction drawn between the ability to discriminate by wavelength alone, and the extent to which a species finds it easy to use such an ability is a useful one. Our knowledge of hearing is far less extensive, but Stebbins has been able to raise a number of interesting questions for further work. The upper limit of pitch for example, varies very markedly between species. The fact that it is relatively low (40–45 kHz) and similar to that of higher primates in the slow moving potto and slow loris, but as high as 60 kHz in the closely related galagos, just as in cat and rat, may give some clues to explain the loss of the ability to hear high sounds which has occurred in most primates.

R. J. ANDREW

KINTSCH, W. *Learning, Memory, and Conceptual Processes.* New York: Wiley and Sons. 1970. Pp. 498. £4·40.

There has long been a need for an undergraduate textbook covering the area of human cognition and learning. Since Kintsch's book goes some way towards filling this gap it should be welcomed despite the shortcomings that must inevitably affect any attempt to survey such a broad and rapidly changing area.

The book begins with an excellent survey of the techniques and theoretical issues of classical verbal learning. This is followed by a chapter on the stimulus sampling theory of Estes, two good chapters on memory, chapters on discrimination learning, concept identification and a final chapter on rule learning and language.

The material is well-organized and the writing clear, although the book does not make particularly exciting reading. This is no doubt partly because the theoretical approaches which clearly appealed to Kintsch at the time of writing are those based on Estes' stimulus sampling theory, and Chomsky's transformational grammar. Stimulus sampling theory has subsequently proved difficult to apply to any but the almost trivially simplified tasks on which it was initially based and has consequently lost most of its theoretical appeal. Its prominent place in Kintsch's text is particularly unfortunate since it is an approach which all but the most mathematically oriented undergraduate is likely to find rather tedious. This is not the case with Kintsch's extensive coverage of the attempt to apply transformational grammar to memory tasks. Unfortunately, however, since the book was written, it has become clear that most if not all the effects on memory previously attributed to syntax are more readily expained in terms of uncontrolled semantic factors that cannot be adequately handled by Chomsky's transformational grammar.

However, although the theoretical ideas underlying much of the book have already been bypassed, it remains a good bread-and-butter text, full of clearly presented information across a wide range of subjects. As such it provides a useful, and by current standards, reasonably-priced background text for courses on human memory and cognition.

<div align="right">A. D. BADDELEY</div>

MORTON, J. (Ed.) *Biological and Social Factors in Psycholinguistics.* London: Logos Press. 1971. Pp. 215. £3·50.

Psycholinguistics is in a mess. Yet only a few years ago it seemed on the threshold of some remarkable discoveries. Its major impetus was Chomsky's transformational grammar and its theoretical consequences for adult linguistic performance. But Chomsky's work also had dramatic implications for the child's acquisition of language. He postulated a distinctive innate component, presumably affecting cerebral organization, that enabled the child to "tune in" to the particular language of his culture. How else could one explain so rapid a learning of so complicated a skill? This revival of "nativism" was carried through in an almost equally rapid time; and soon theorists were arguing that the differences between languages were "superficial"—their underlying "base" component was universal. In short, it seemed as though both the adult's linguistic performance and the child's acquisition of langugage would be illuminated by transformational grammar. What went wrong? The present book tells most of the story.

The book evolved out of a symposium convened by John Morton at the XIXth International Congress of Psychology held in London during the summer of 1969. It consists of nine papers, of which only four were actually presented at the Congress, and between them they bring the history of psycholinguistics up to date.

The first thing to go wrong was that experiments on adult performance ceased to come up with the right answers. This part of the story has been told before, but the paper by Tom Bever, which takes up nearly a quarter of the book, retells it with exemplary clarity. Moreover, Bever has the happy knack of turning even the current confusion in linguistics to his advantage. Who else would have discovered that competing accounts of syntax can account for the recall of sentences at different times after their original presentation?

The second thing to go wrong was that an experiment that had always failed in the past was finally made to work. The Gardners succeeded in teaching a chimpanzee to talk: an

animal called Washoe has evidently mastered a large part of the vocabulary of a gestural sign language, and is able to combine signs to convey simple messages. This startling achievement, which other chimpanzees have subsequently emulated, naturally calls for a reconsideration of the nativist position. Just how important is man's biological make-up to the development of language?

Of course, it all depends upon what is meant by "language" and this is one of the central issues of the book. Bernard Campbell takes it up in the opening chapter during the course of his survey of the biological background to language. His view is simply that human language consists of the ability to name, combined with the older and more primitive skill of communicating emotion. However, as John Marshall points out in the very next chapter, there is a controversy about naming. Sooner or later a child has the insight that it is possible to talk about anything. This is characterized sometimes as the insight that everything has a name, and sometimes as the insight that everything can have a predicate applied to it. Unfortunately, there seems to be no decisive evidence either way. Hence, in tracing the stages of the child's linguistic development, Marshall suggests the "real" use of language comes with the ability to combine separately learnt signs in a semantically appropriate manner, and with the ability to discriminate between the correct and incorrect use of such combinatorial symbols. It looks very much as though Washoe and her colleagues are going to pass these tests with flying colours.

The whole venture is rendered still more paradoxical by the findings which I. M. Schlesinger reports in his chapter. It appears that human users of an Israeli sign language for the deaf have no consistent rules for signalling such grammatical relations as the subject and object of the sentence, though they do have rules governing the position of the verb. Thus, it begins to look as though no simple syntactic distinction can be drawn between the human being's and the chimpanzee's use of sign language. However, there may be other criteria which will prove decisive for human language. The editor of the book in his own contribution stresses the importance of "displaced" speech, i.e. speech concerning objects or events not actually present before the participants. Ruqaiya Hasan points out that a competent speaker must have rules governing the appropriateness of discourse to its context; and John Marshall goes still further in postulating the speaker's need, not only for an internal representation of context, but also for a representation of the listener's potential states of mind. Such, at least, seems to be necessary in order to account for a speaker's ability to deceive the listener. Does Washoe, one wonders, tell lies? The answer might tell us as much about her mental life as her morals.

If there is so much uncertainty about the nature of "real" language, it is hardly surprising that the status of linguistic universals and the innate component should be highly controversial. Schlesinger's results, of course, cast doubt on the universality of the base component—if one is prepared to accept that he is working with a "real" language. John Morton questions whether the innate component is especially important for language, and even whether it is specific to language. He makes the very good point that claims about the complexity of language and the speed of its acquisition exist in an intellectual vacuum. No other department of intelligent human activity has been sufficiently elaborated for a meaningful comparison. Similarly, Roger Wales points out that a number of the child's linguistic skills seem to rest upon mechanisms developed for perception. The ability to handle relational adjectives, for example, is analogous to being able to transfer a perceptual relation from one range of stimuli to another. This argues for a greater continuity with other species, since it is known that they can master the perceptual task.

The possibility that perceptual or cognitive mechanisms may underlie linguistic universals is discussed by David McNeil, in a short note on the topic. He contrasts this idea with the concept of "strong" linguistic universals which have no such origin. It seems, for instance, that some strictly linguistic property is required to make a word into a noun. The fact that anything can be named—actions, events, entities, etc.—is a principle which cannot be accounted for on purely cognitive grounds. Once again, however, a controversy seems to be in the offing since John Morton claims that naming involves merely an extension of such operations as generalization and discrimination which one finds in lower species.

One of the strengths of the book is that it concentrates upon a small number of specific topics. But it would be a mistake to leave the reader with the impression that it is solely about the nature of "real" language and the status of linguistic universals. The wide-ranging nature of Bever's contribution has already been mentioned; and it should also be noted that Hasan provides a similarly broad review of the main problems of relating syntax to semantics. The distinguished Soviet psychologist, A. A. Leontiev, enters a modest plea for the role of social factors in the development of language. Finally, in the midst of all the theoretical speculation and reappraisal, it is refreshing to read Roger Wales's account of his sustained and ingenious experimental attacks upon a number of basic psycholinguistic problems.

P. N. JOHNSON-LAIRD

RITCHIE RUSSELL, W. *The Traumatic Amnesias.* London: Oxford University Press. Pp. 84. £2.00.

Ritchie Russell has contributed extensively to our knowledge of traumatic amnesia. This book contains selections from nineteen of his books and papers written between 1932 and 1968 together with a very limited amount of connecting prose. The subject matter ranges from discussions of the nature of traumatic brain injury to theories of memory. To a psychologist the first chapter which deals with the physical consequences of traumatic brain injury is of background interest only. The central section of the book, some 50 pages, is a natural history of the effects of traumatic injury on memory. These core papers were of considerable importance when written, but they suffer severely when collected in book form, as there is virtually no reference to other work and naturally none to subsequent work. The brief connecting text does not remedy this defect.

Certain aspects of the natural history are probably little known to psychologists. Visions can occur that are obviously related to the injury, as they depict events which are known from other evidence to have been part of the accident. Yet they are not consciously associated with the accident by the patient. Also retrograde amnesia tends to be more likely to occur with left hemisphere damage than with right. Unfortunately this section tends to be somewhat uncritical. For instance, no mention is made of the possibility that the measured "duration" of retrograde amnesia, other than that of a few seconds, could be an artefact resulting from a general retrieval failure which is most noticeable for events just before the accident, since these are what the patient is asked about. When consciousness is regained it may well be difficult to provide appropriate recall cues for these somewhat distant and, at the time, possibly trivial events, which will not be connected with later events.

The final two chapters on the nature of memory are of little value as they present relatively widely held speculations without alternative theories being considered. In this section, particularly, the method of republishing old papers is rather pointless as theories in psychology age more rapidly than does the natural history type of empirical work.

TIM SHALLICE

MILLER, N. E. *Selected Papers.* Chicago: Aldine Publishing. 1971. Pp. 874. £15.00

Neal Miller is certainly an eminent figure in experimental and physiological psychology, and no collection of the papers of any other person working in this field could be as welcome. Within this selection is illustrated the wide range of interest and the impressive quality and quantity of his work. It is possible that nobody has been more influential in the rapid recent advance of physiological psychology than Miller, for it is work from his laboratory and under his direction that has in many cases clarified the issues, and turned initial experimental observations into well formulated and coherent areas of investigation.

What picture then does this collection of papers give of Miller's approach? First, there appears a down-to-earth common sense in experimenting on the brain and behaviour. We find a series of no-nonsense empirical enquiries, well designed, addressing clearly conceived problems, and with the results expounded in superbly written review articles with due regard for both the particular and the wider implications. Second, Miller's work is

characterized simultaneously by dogged determination to pursue certain problems and an ability to take up relevant new lines of investigation. Thus one of the main themes of his work has been on the notion of fear of pain as an acquired drive. Enquiring whether animals would learn to avoid onset of other supposedly aversive states such as hunger or thirst was a logical next question, which Miller hammered at (unsuccessfully) for many years, and it was in this context, he says, that electrical brain stimulation as a means of rapidly initiating hunger or thirst acquired interest as a technique which might allow him to demonstrate the effect. As a matter of fact it did not, but Miller was undeterred and tried yet other methods to approach this same problem, while becoming more and more involved in direct manipulation of the brain as a means for inducing and reducing drives. His method has been to proceed purposefully towards clearly defined objectives, but at the same time approaching with a sufficiently wide programme of experimentation for there always to be promising developments that could be capitalized upon. Third, and perhaps most characteristically, Miller has been enormously successful in inspiring, and guiding graduate students. All but a small handful of Miller's experimental papers are collaborative, and most of them report research done with graduate students, and result from the working out of the themes that Miller has instigated and continuously developed during the 35 years which this book covers.

The volume contains 64 papers that Miller himself has chosen from a much larger total on the grounds of historical relevance and current interest. They are grouped into nine sections illustrating the various themes he has followed. There are sections on the effects of drugs upon behaviour; on learning and secondary reinforcement; on the stimulus response formulation and drive reduction; on hunger and thirst, and on the motivational effects of chemical and electrical stimulation of the brain. Lastly, comes his most recent work on instrumental conditioning of autonomic responses such as heart rate, blood pressure, urine flow and so forth. With this he returns again to problems of possible clinical and applied significance which marked what is perhaps his best known and most influential research on conflict and fear, which is represented as the opening sections of the volume.

This book certainly deserves a place in every library of psychology. The work which is included in it has not changed the direction of psychology, but instead represents the best type of experimental investigation within the context of what Kuhn would call normal science. Re-reading some of these papers, the idea occurs to one that perhaps the contribution of such sustained consolidative efforts as Miller's is just as important for progress in science as that of revolutionary innovations.

KEITH OATLEY

SUTHERLAND, N. S. AND MACKINTOSH, N. J. *Mechanisms of Animal Discrimination Learning.* New York: Academic Press. 1971. Pp. 559. $18.50.

This is without question a very good book: well thought-out, well argued, and well written. It is not, however, a book for the casual nibbler. What the reader is offered is meat of the highest quality, readily digestible, but requiring a degree of application.

Against a rough sketch of the history of selective attention, the authors present their own theory of the mechanisms primarily responsible for attention-like phenomena in animals. There follows a foray through many of the classical phenomena of animal learning: continuity and non-continuity, generalization, reversal learning, reversal and non-reversal shifts, serial reversal learning, partial reinforcement and extinction, probability learning etc.

The ability of the authors' theory to handle such diverse phenomena is remarkable, and must be accounted among the major achievements in the field of animal learning. At a time when knowledge of behavioural phenomena has far outstripped our ability to explain them, it is heartening to see a bold theoretical approach backed by persistent experiment. The theory is somewhat loosely formulated, and is not completely successful in several particulars. It is clear that a complete explanation of many of the phenomena discussed will incorporate concepts of frustration and memory, which are not included in the authors' theory, though some are discussed in the book.

The problem, as the authors see it, is that "in theorizing about animal learning, workers are faced with a choice between putting up formalized models to explain a narrow range of data and working over a much broader range of phenomena with much looser models". But this is not a problem peculiar to psychology. It is common in many sciences, where it is met by a certain degree of sophistication concerning the nature of models and their relation to observed phenomena. In view of the enormity of the problem of explaining behaviour, the present day battery of theoretical apparatus seems hopelessly naive. Almost alone amongst scientists, psychologists take little account of modelling theory in general, or of well-established explanatory concepts.

Revulsion against the pathetic system-building of the thirties and forties seems to have led to an entrenched phenomenalism. The authors of this book are to be congratulated on having steered a steady course across this current of opinion, and on making a spectacular landfall, perhaps on an island, but hopefully on a continent.

D. J. McFarland

FURCHTGOTT, E. (Ed.). *Pharmacological and Biophysical Agents and Behavior*. New York: Academic Press. 1971. Pp. xii+402. $14.50.

The agents reviewed are ionizing radiations (Furchtgott), microwave irradiation (Thompson and Bourgeois), hypoxia (Meier), ambient temperature (Pepler) and several classes of drugs—convulsants, xanthines and pemoline (Calhoun), atropine and scopolamine (Carlton and Markiewicz) and sympathomimetic amines (Grossman and Sclafani). Most chapters have a brief physiological introduction and then cover the literature on behavioural effects of these agents, fairly critically and coherently for the space used. The chapter on anticholinergics is more selective and goes into detailed behavioural interpretation. Most of these topics have been reviewed recently, but only as separate items scattered in more general volumes. There are now some psychopharmacology textbooks but this collection provides a more advanced and detailed treatment of the topics selected.

D. A. Booth

EYSENCK, H. J., ARNOLD, W. AND MEILI, R. (Eds). *Encyclopedia of Psychology*. *Volume 1*. London: Search Press. 1972. Pp. 396. £8.00.

This is the first of three volumes in which dictionary definitions of psychological and psycho-analytical terms are assembled together with nearly 300 longer entries on topics of major interest and a number of biographical notes on eminent psychologists. The total number of entries for the three volumes is approximately 5000. Contributions are drawn from 22 different countries and the work is to be published in English, German, French, Spanish, Portuguese and Italian. Volume 1 deals with items A–F. The remaining two volumes will follow at a later date. It is an ambitious venture and although the three volume set will cost £24, there could be a fair demand for a work which *The Times* reviewed under the heading, "All you ever wanted to know about Psychology".

How should such a work be assessed? It should presumably be convenient to use, contain a comprehensive selection of terms, be representative and should fill a gap not met by other publications.

Unfortunately, the encyclopedia is handicapped from the start by a bad choice of format. It is printed (in double columns) in large print on thick paper with a considerable amount of white space between entries. In consequence, the number of items in the set of three volumes does not greatly exceed the number in the pocket size *Penguin Dictionary of Psychology* (1952, 1964), barely reaches the number in Warren's dictionary (1934) and is less than half those dealt with in the compact single volume dictionary of English and English (1958).

Clearly then, one will find many items in the older dictionaries which are not dealt with in this new encyclopedia. Some terms may have been excluded because they are no longer in current use. *Aufgabe* may be only of historical interest. The demise of Hullian learning theory may make it unnecessary to look up *anticipatory goal response,* and *cell assemblies* are perhaps no longer fashionable. However, some of the choices still look

rather arbitrary. One looks in vain for *cochlea* although *chiasm, optic* is there. *Corpus callosum* is missing, although *corpus luteum, c. quadrigemina,* and *c. striatum* appear. For *epilepsy* one is referred to *abence,* which does not appear as an entry and although *aphasia* gets a full article, *amnesia* and *amnesic syndrome* are each given only short entries. The more physiologically inclined reader may regret the passing of *Brodmann's areas* and the *Bunsen-Roscoe law.* One is also occasionally surprised by the inclusion of terms used in a non-technical sense such as *dating, diagram, diction, film* ("an audio visual mass medium . . ." etc.) and *cinema* (cross reference to *film*).

The work is, of course, more than a dictionary and many of the longer articles (which include useful bibliographies) well repay reading. The English reader, however, may not find quite what he expects and when one examines the list of contributors the reason becomes apparent. Although contributors are drawn from 22 countries, nearly half the articles in the three volumes are written by contributors from West Germany. This gives many of them a flavour distinctly different from the normal diet of the British university student, who may enjoy having his palate tickled by statements such as "the tasks proper to applied psychology focus in general on the determination of character traits" (p. 76).

The short biographies form a very small part of the content. There are only 15 such entries in the first volume (none of them referring to psychologists now living) and no separate list of the names is given—at least, not in Volume 1. In most cases one could obtain more information more quickly from other sources.

Poor proof-reading of some of the mathematical formulae has led to odd results. H (average uncertainty) suffers particularly badly, both under *binary digit* and under *communication.*

The encyclopedia therefore has its limitations, both in convenience and in comprehensiveness. The price will probably deter individuals who wish to buy a reference work for their own collection, but the set should certainly be acquired by most libraries dealing with psychology. It seems unlikely that it will replace the more comprehensive reference works which one finds in larger libraries, but it should find a useful place among them. Perhaps its greatest use will be in small libraries of a few thousand volumes and in this role it deserves a welcome, although two cheers may be more appropriate than three.

R. DAVIS

References

ENGLISH, H. B. and ENGLISH, A. C. (1958). *A Comprehensive Dictionary of Psychological and Psychoanalytic Terms.* New York, London, Toronto: Longmans & Green & Co.

DREVER, J. (1952). *A Dictionary of Psychology.* Harmondsworth: Penguin. (Revised edition, 1964, by Wallerstein, H. and Singer, B. R.)

WARREN, H. C. (1934). *Dictionary of Psychology.* Cambridge, Massachusetts: Houghton Mifflin.

Publications Received

ARONSON, L. R., TOBACH, E., ROSENBLATT, J. S. and LEHRMAN, D. S. (Eds). *Selected Writings of T. C. Schneirla.* San Francisco: W. H. Freeman & Co. 1972. Pp. 1032 + xiv. £10·20.

BAKKER, D. J. *Temporal Order in Disturbed Reading.* Rotterdam University Press. 1972 Pp. 100. Dfl. 35.

BARTLEY, S. H. *Perception in Everyday Life.* New York: Harper & Row. 1972. Pp. 325. £1·50.

BOAKES, R. A. AND HALLIDAY, M. S. *Inhibition and Learning.* London: Academic Press. 1972. Pp. 568. £7·25.

BROZEK, J. AND SLOBIN, D. I. (Eds). *Psychology in the USSR: an Historical Perspective.* New York: International Arts and Sciences Press. 1972. Pp. 301. $15.00

DARTNALL, H. J. A. (Ed.). *Handbook of Sensory Physiology.* Berlin: Springer-Verlag. 1972. Pp. 810. $59.30.

DENENBERG, V. H. (Ed.). *Readings in the Development of Behavior.* Reading: W. H. Freeman & Co. Ltd. 1972. Pp. 483. £2·90.

DODWELL, P. C. *Perpetual Processing.* New York: Appleton-Century-Crofts. 1972. Pp. 535. £16·50.

DUBRIN, A. J. *The Practice of Managerial Psychology.* Oxford: Pergamon Press Ltd.: 1972. Pp. 326. £4·50.

EKMAN, P., FRIESEN, W. V. AND ELLSWORTH, P. *Emotion in the Human Face.* Oxford: Pergamon Press Ltd. 1972. Pp. 191. £3·50.

GEDDES, L. A. *Electrodes and the Measurement of Bioelectric Events.* New York: Wiley Interscience. 1972. Pp. 364. £6·75.

GLASS, D. C. AND SINGER, J. E. *Urban Stress.* New York: Academic Press. 1972. Pp. 182. $8.75.

GRAHAM, D. *Moral Learning and Development: Theory and Research.* London: Batsford. 1972. Pp. 317. £3·00.

HARRISON, B. *Meaning and Structure: an Essay in the Philosophy of Language.* London: Harper & Row Ltd. 1972. Pp. 318. £4·65.

HINDE, R. A. (Ed.). *Non-verbal Communication.* Cambridge: Cambridge University Press. 1972. Pp. 443. £5·00.

JOHNSON, H. H. AND SOLSO, R. L. *An Introduction to Experimental Design in Psychology: A Case Approach.* London: Harper & Row Ltd. 1972. Pp. 216. £1·85.

JORDAN, N. *Themes in Speculative Psychology.* London: Tavistock Publications Ltd. 1968. Pp. 259. £0·75.

KOLERS, P. A. *Aspects of Motion Perception.* Oxford: Pergamon Press Ltd. 1972. Pp. 220. £4·00.

LEVINE, G. AND BURKE, C. J. *Mathematical Model Techniques for Learning Theories.* New York: Academic Press. 1972. Pp. 288. $12.50.

McCULLOCH, J. W. AND PHILIP, A. E. *Suicidal Behaviour.* Oxford: Pergamon Press. 1972. Pp. 123. £2·50.

McGAUGH, J. L. *Psychobiology: Behavior from a Biological Perspective.* London: Academic Press. 1971. Pp. 366. £6·75.

MOLTZ, H. *The Ontogeny of Vertebrate Behavior.* New York: Academic Press. 1971. Pp. 500. $22.00.

MUSSEN, P. H. AND ROSENZWEIG, M. R. (Eds). *Annual Review of Psychology.* Palo Alto, California: Annual Reviews Inc. 1972. Pp. 786. $10.50.

NEWBOLD, H. L. *The Psychiatric Programming of People.* Oxford: Pergamon Press Ltd. 1972. Pp. 110. £3·00.

NEWELL, A. AND SIMON, H. *Human Problem Solving.* Englewood Cliffs: Prentice-Hall Inc. 1972. Pp. 920. $16.95.

NEWTON, G. AND RIESEN, A. H. (Eds). *Advances in Psychobiology.* Vol. 1. New York: Wiley Interscience. 1972. Pp. 302. £6·50.

OLSON, D. R. *Cognitive Development.* London: Academic Press. 1971. Pp. 220. £5·60.

PEARSON, R. *The Avian Brain.* London: Academic Press. 1972. Pp. 658. £10.

PIAGET, J. *The Principles of Genetic Epistemology.* London: Routledge & Kegan Paul. 1972. Pp. 98. £1·50.

PRIBRAM, K. H. *Languages of the Brain.* Englewood Cliffs, N.J.: Prentice-Hall International. 1971. Pp. 432. £4·75.

PRINGLE, J. W. S. (Ed) *Biology and the Human Sciences.* Oxford: Clarendon Press. 1972. Pp. 139. £2·50.

RUNES, D. *Handbook of Reason.* New York: Philosophical Library Inc. 1972. Pp. 200. $6.00.

SARASON, I. G. *Personality: an Objective Approach.* Second Edition. New York: John Wiley & Sons. 1972. Pp. 601. £4·55.

SCHNEIRLA, T. C. *Army Ants: a Study in Social Organisation.* Reading: W. H. Freeman & Co. 1972. Pp. 349 + xx. £5·70.

SIEGMAN, A. W. AND POPE, B. (Eds). *Studies in Dyadic Communication.* Oxford: Pergamon Press Ltd. 1972. Pp. 336. £5·50.

SLUCKIN, W. *Imprinting and Early Learning.* London: Methuen & Co. Ltd. 1972. Pp. 180. £2·00.

TULVING, E. AND DONALDSON, W. (Eds). *Organization of Memory.* New York: Academic Press. 1972. Pp. 423. $17.50.

YOUNG, T. R. *New Sources of Self.* Oxford: Pergamon Press Ltd. 1972. Pp. 114. £2·75.

ZEIGARNIK, B. V. *Experimental Abnormal Psychology.* New York: Plenum Publishing Corp. 1972. Pp. 158. $20.

Notes for Authors

The Journal publishes original papers on experimental work in all branches of human and animal psychology. Papers dealing with the mechanisms underlying behaviour are particularly welcome and theoretical papers on such topics are acceptable provided they relate to experimental findings. Short notes on apparatus likely to be of interest to experimental psychologists are also acceptable, provided that the design of the apparatus is an original contribution by the author.

Papers should be submitted in duplicate, one copy to be a *typewritten original*, preferably on A4 paper (219 × 297 mm), in double spacing with wide margins.

Title The title should be as concise as possible and should appear as a separate sheet together with the name(s) of the author(s) and the name of his institution. (The full postal address is not required on the title page.) A short running title of not more than 40 characters (including spaces) should be indicated if the full title is longer than this.

Abstract An abstract of 50–200 words should follow the title page as a separate sheet.

Headings and sub-headings for different sections of the paper should be clearly indicated. Authors should follow the conventions used in current issues of the Journal.

Acknowledgements should be as brief as possible and should appear as a separate sheet at the end of the text.

Footnotes should be avoided unless absolutely necessary. Essential footnotes should be indicated by superscript figures in the text and collected as a single page at the end of the manuscript.

References in the text should be indicated with the year of publication, in parentheses after the author's name, If more than one article by the same author in the same year is quoted, *a, b*, etc., should follow the year. When a paper has been published jointly by two authors, both names should be given in the text. For more than two authors, quote all names at the first mention and use the first author '*et al.*' for subsequent mentions. References to unpublished work should be avoided since they are of little use to readers. A list of the references quoted in the text should be given at the end of the paper in alphabetical order of author's names, commencing as a new sheet. Titles of Journals and books should be given *in full* thus:

WOODWORTH, R. S. (1901). On the voluntary control of the force of movement. *Psychological Review*, **8**, 350–9.

GALTON, F. (1893). *Inquiries into Human Faculty and its Development*. London: Macmillan.

GARRETT, M. and FODOR, J. A. (1968). Psychological theories and linguistic construction. In DIXON, T. R. and HORTON, D. L. (Eds), *Verbal Behavior and General Behavior Theory*. Pp. 451–77. Engelwood-Cliffs, N.J.: Prentice-Hall.

Tables should be kept to the minimum consistent with clairty of exposition.

Each table should be typed double spaced as a separate sheet, giving the heading, e.g. TABLE II, in Roman numerals, followed by the legend, followed by the Table. Make sure that appropriate units are given.

Instructions for placing the Table should be given in parentheses in the text, e.g. (TABLE II about here).

i

Figures are expensive to reproduce and should only be used when essential. The same data should not be presented both as a figure and in a table. Where possible, related diagrams should be grouped together to form a single figure. Figures should be drawn to professional standards in Indian Ink on Bristol Board of A4 size, and it is recommended that the linear dimensions of figures be approximately twice those intended for the final printed version. Lettering to appear in the figure should be done on tracing paper placed over the figure. Make sure that axes of graphs are properly labelled and that appropriate units are given.

The legend should be typed on a separate sheet of paper headed, e.g. FIGURE 2, in Arabic numerals. Instructions for placing the figure should be given in parentheses in the text, e.g. (FIGURE 2 about here).

Statistics   Greek letters should be used to represent population parameters and Roman letters for sample values. Results of statistical tests should be given in the following form:

$$F = 25 \cdot 3, \; df = 1,9, \; P < 0 \cdot 001.$$

Similarly for t, χ^2 and other tests.

Units   The Système Internationale (SI) will be used for all units. (If measures such as inches are used, then SI equivalents should be given.) If in doubt, refer to *Units, Symbols and Abbreviations*, published by the Royal Society of Medicine, 1 Wimpole Street, London, W.1.

The author will normally receive a decision on his paper within 3 months of receipt and if it is accepted it will normally be published 3 to 6 months later. The date of the receipt of the manuscript will be printed. Where minor revision of a paper is requested the original date of receipt will appear, provided that a satisfactory revision is received within 1 month of the request.

Contributors receive 25 copies of their paper free. Additional copies may be ordered on a form provided by the publishers at the time proofs are sent to authors.

OCCUPATIONAL PSYCHOLOGY

Editor: ISABEL BLAIN

Contents

CONTENTS

629/
9.4.73

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY

THE QUARTERLY
JOURNAL OF
EXPERIMENTAL
PSYCHOLOGY

Volume 24 1972

Printed in Great Britain

Volume 24

Number 1, February 1972

iii

CONTENTS

NUMBER 4, NOVEMBER 1972

RICHARD CHARLES OLDFIELD

OBITUARY NOTICE

RICHARD CHARLES OLDFIELD (1909-1972)

Carolus Oldfield, as he was always known, died suddenly on 27 April 1972, aged 62. He was a founder member of the Experimental Psychology Group (the precursor of the present Society) and its President from 1956–1957. He was the first Editor of this *Journal*.

Richard Charles Oldfield was educated at Marlborough College and Peterhouse, Cambridge, where he read Natural Sciences followed by Moral Sciences (as Philosophy was then known at Cambridge) and after graduating was awarded the Arnold Gerstenberg Studentship for research in the philosophy of science. But his interests soon took him into experimental psychology and up to the outbreak of war he worked in the Cambridge Psychological Laboratory under Sir Frederic Bartlett. After the war, in which he served with distinction in the Radar Branch of the R.A.F.V.R., Oldfield was appointed to a University Lectureship at Oxford, where he played a key role in the establishment of the Honours School of Psychology, Philosophy and Physiology (P.P.P.) and in fostering research at the Institute of Experimental Psychology. He later held the Chairs of Psychology at the University of Reading (1950–56) and Oxford (1956–66). In 1963, he was appointed Honorary Director of the M.R.C. Psycholinguistics Unit and, three years later, on the reconstitution of the Unit as the Speech and Communication Research Unit and its transfer to a new base in Edinburgh, Oldfield became its full-time Director. At the time of his death, he was an Honorary Professor in the University of Edinburgh.

Oldfield's early work lay in what is now called philosophical psychology but his unpublished dissertation on "Reasoning, Language and Logic" illustrates well the transition of his interest to issues in experimental psychology. Thereafter, all his work fell within this sphere. At Cambridge, he published several papers on learning and memory, together with a critical review (jointly with the present writer) of Head's concept of the Schema and the use made of it by Bartlett in his book on *Remembering*. He also drew together the outcome of a number of lines of work then proceeding at the Cambridge Psychological Laboratory in a noteworthy paper entitled "Some recent experiments bearing on 'internal inhibition' ", read at a joint meeting of the British Psychological Society and the Cambridge Psychological Society in November 1936. At this time, too, Oldfield developed a lively interest in the special senses, particularly taste and smell, though in this field he published little.

In 1936, Oldfield was awarded a Leverhulme Research Fellowship to study techniques of interviewing under the aegis of C. S. Myers. The outcome of this was a lively and provocative little book on *The Psychology of the Interview*, published

in 1941. Although the book was criticized in some quarters for its lack of quantitative treatment, Oldfield was too good a scientist to apply quantitative procedures to issues in which they were doubtfully applicable and almost certainly unhelpful. Quantitative treatment, he maintained, was beyond all doubt the aim of experimental psychology but if only for this reason should be embarked upon with appropriate circumspection. The use of numbers merely to enhance scientific prestige was to him totally abhorrent.

After the war, Oldfield returned for a time to problems in sensory psychology, publishing several papers on psycho-acoustic issues. Of these, the most noteworthy was a study of apparent fluctuations of sensory thresholds based on work begun at Oxford but completed at Reading and published in this *Journal* in 1955. He also supervised the post-graduate work at Oxford of R. B. Joynson and J. Langdon on problems of perceptual constancy, which likewise appeared in this *Journal*. He became interested in the perceptual problems presented by radar, films and television, contributing—along with Professor A. Michotte—to a symposium on this topic published in a French periodical in 1948. Some years later he produced for the *Annual Review of Psychology* an admirable review article on somaesthesis.

But with passage of time, Oldfield's early interest in the psychology of language gradually re-asserted itself. This was no doubt in part due to the renascence of linguistic interests among psychologists generally, catalysed at Oxford by George Miller during a period of sabbatical leave. It also owed much to the developing techniques of psycho-acoustic inquiry, in particular dichotic listening, which were being developed at Oxford during his tenure of the Professorship by a number of gifted research students, among them Anne Treisman and Neville Moray. A further strand in this line of advance was Oldfield's long-standing interest in aphasia and kindred disorders of speech—an interest which could be dated from a period early in the last war when he worked for some months in Sir Hugh Cairns' Department of Neurological Surgery at Oxford. This interest led to important post-war links with the Oxford Department of Neurology and to collaborative work with Margaret Elvin, Moyra Williams and later Freda Newcombe, on the study and analysis of language disorder in patients who had sustained penetrating brain-wounds. Although Oldfield's name appeared on only a few of the resulting publications, of his stimulation, interest and advice there can be no question. Much of the work, both experimental and clinical, carried out at Oxford by the Psycholinguistics Unit was admirably summarized by Oldfield in his Sir Frederic Bartlett lecture—the first in the series—published in this *Journal* in 1966 under the title of "Things, Words and the Brain".

It was, however, in his final years at Edinburgh that Oldfield, now largely freed of administrative and teaching duties, was perhaps at his most creative. With John Marshall, he published an admirable volume of readings on the psychology of language. He re-opened some old questions on handedness and its relation to hemisphere dominance, presenting an excellent paper on this topic at the 9th International Congress of Neurology in New York in September, 1969. He wrote interesting and stimulating papers on handedness in musicians and on the psychology of music. One of his last and most challenging papers was entitled "Some

impediments to verbal communication", read at a symposium of the Institute of Biology in March, 1971, not yet published but fortunately available as a Report from the Speech and Communication Unit. In this paper, Oldfield faced squarely the implications of new styles in communication and social interaction for several basic issues in psychology. It is a matter of deep regret that he did not survive to carry these preoccupations further.

Oldfield's loyalty to experimental psychology was unchallenged. While appreciating its limitations and the excessive optimism of its pioneers, he had no doubt that such advances as might be achieved in psychology would come about only through the use of experimental method geared to theory of sufficient power to generate worthwhile experiments. He had no use for facile scientism or for premature claims on behalf of psychology envisaged as a professional subject. As he saw it, psychology would be much better advised to attract the interest of able young men and women trained in the various fields of natural science than to set itself up as an exclusive and self-contained academic discipline. He was thus to some extent out of key with post-war developments in the subject, which perhaps explains the fact that, while honoured by psychological societies abroad, he received little honour in his own country. Nonetheless, experimental psychologists in Britain will wish to pay tribute to the memory of a charming, accomplished and devoted colleague, to whom many of us owe much, both intellectually and in terms of personal friendship.

O. L. Z.

REHEARSAL STRATEGY AS A FUNCTION OF RECALL EXPECTATION

LAIRD S. CERMAK†

Psychology Service, Boston Veterans Administration Hospital, Massachusetts

The subjects were divided into three groups with respect to their expectations concerning a recall task given after the final trial of the usual STM distractor procedure. Group 1 were told only that they would have to recall during each trial's recall interval, thus did not expect to have to recall again. Group 2 were told that they would have to recall all the words presented in the experiment at the end of the last trial, in addition to the trial-by-trial recall. Group 3 were told only that they would have to recall after all words had been presented and they sat passively through the presentation trials. In addition to their recall expectations, half of the subjects in each group received a 2-s presentation and half received a 5-s presentation interval. It was found that the length of the presentation interval had an effect on the number of words recalled at the end of all trials, but recall expectancy did not. However, expectancy did determine the rehearsal strategies of subjects and hence the serial positions from which items were recalled.

Introduction

The Brown–Peterson distractor technique has frequently been used in the study of encoding and short-term memory (STM). The most widely reported finding is that the probability of recall decreases from the first to the fourth or fifth recall trial. One possible explanation for this decline in performance is that the subject's encoding efficiency deteriorates as the experiment progresses. The alternative to this explanation is that the encoding efficiency remains constant but the number of residual traces available to interfere with retrieval increases as the number of presentations increases.

The latter alternative, taken at face value, would imply a linear relationship between the number of presentations and the amount of interference. However, Peterson and Gentile (1965), Loess and Waugh (1967) and Cermak (1970) have all shown that interference produced by the traces of preceding items diminishes as a function of the length of time between recall intervals. This finding suggests two possibilities with respect to the fate of a trace after its initial formation: either the trace simply decays as a function of time or the time between trials allows the traces to be more easily discriminated from one another at recall.

Given the two alternative explanations for the manner in which items are initially encoded, and the two possibilities concerning the fate of a trace over time, three predictions can be made about the relative strength of each trace at the con-

†Full address: Psychology Service, Boston Veterans Administration Hospital, 150 South Huntingdon Avenue, Boston, Massachusetts 02130, U.S.A.

378

clusion of a distractor technique experiment: first, if the declining efficiency in encoding assumption is correct, it would be expected that traces formed on initial trials would be stronger than later trial traces (1). Second, if all items are equally well encoded and their traces do not decay across time, then all traces would be equally strong (2). Third, if all items are equally well encoded but their traces decay as a function of time, then the initial trial traces would be weaker than later trial traces (3).

The present report proposes to ascertain which of these assumptions about encoding is correct by giving a free recall task at the conclusion of the final trial of a distractor type experiment. In this task the subject would be asked to recall as many of the words presented in the entire experiment as possible. If a primacy effect occurs, then alternative (1) above is supported; a recency effect would support alternative (3); and, if recall is evenly distributed across presentation trials, alternative (2) would receive support.

Previous hypotheses about strength of encoding (Wickens, 1970) and decay of trace strength (Cermak, 1970) favour the outcome predicted by alternative (3); i.e. a recency effect. However, if this prediction is fulfilled, it does not automatically follow that memory traces would decay in a situation in which the subject knew they would have to be retrieved a second time. Perhaps decay occurs only in cases in which the subject sees no value in keeping the traces active. Therefore, it was decided to investigate whether retrieval at the conclusion of the experiment is the same for subjects expecting the post-test recall task as it is for those not expecting the task. In other words, if an item's representation in memory does decay following its initial retrieval, is it possible for the subject to prevent this decay.

Finally, this experiment also compared the effect of using a long (5 s) presentation interval with that of a short (2 s) interval. The necessity for this comparison was pointed out by Cermak and Levine (1971) who suggested that more semantic encoding takes place during a 5-s presentation interval than during a 2-s presentation interval and that encoding during the 2-s presentation interval is limited primarily to an acoustic level. If this is true then, as Baddeley and Dale (1966) and Adams (1967) have suggested, information presented for 5 s should attain a more permanent state (in LTS), and thus should be more available during post-test recall, than information presented for only 2 s.

Method

Subjects

The subjects were 180 undergraduate psychology students at Tufts University who chose to participate in this experiment as partial fulfilment of a course requirement. They were assigned to one of the six groups in order of their appearance.

Design

The Peterson and Peterson (1959) version of the distractor technique was used. In this procedure a word triad was projected on a screen for a predetermined amount of time; the subject was then distracted from rehearsal until a signal for recall was presented. The distractor task consisted of counting backward by threes from a number flashed on a screen, in time with a metronome beating at one beat per s. Additional trials followed consisting of

different word triads but using the same procedure. Six sets of word triads were presented consisting of 18 common English words, bearing no obvious acoustic or semantic relationship to each other.

Procedure

The groups were divided on the basis of the length of the presentation interval: one-half of the subjects received a 2-s visual presentation interval and one-half received a 5-s visual presentation interval. Within each of these groups the subjects were further subdivided on the basis of the recall instructions they received: one-third of the subjects were given the usual distractor technique recall instructions, i.e. they were informed only that they would have to recall during the initial trial-by-trial recall intervals (Initial only). Another third were told that they would not only have to recall on each trial, but would later be asked to write down all the words they had seen at the conclusion of the experiment (Initial plus Post-test). The final third were told to recall only after all trials had been concluded (Post-test only). In other words this last group did nothing overt during the usual recall interval. These conditions resulted in a total of six groups in the experiment.

The following sequence of events and timing was presented for each trial: an asterisk (ready signal), 2 s; word triad presentation, either 2 or 5 s; retention interval, 15 s; recall interval, 10 s; and a rest interval of 6 s. The subject was presented with six trials, and his responses during each recall interval were recorded. Immediately following the final trial, all subjects were asked to write down as many of the words that had been presented as they could.

Apparatus and materials

The subject was seated in a small room, facing a screen, upon which the stimulus material was projected by a Kodak Carousel 800 projector. The timing of events was programmed by a Gerbrands tape timer. The only other apparatus employed was a metronome beating at one beat per s (used to pace distractor task). Four cards were taped to the wall to acquaint the subject with the experiment.

The words used in this study were all common (Thorndike–Lorge A count words) English words of 4–6 letter length. Care was taken to insure that no synonyms or homonyms were included in the list and that no two words in the same triad began with the same letter. These word triads were completely counterbalanced across the six trial positions.

Results

Initial Recall

The percentage of total words that were correctly recalled during the initial retrieval interval across all six trials resulted in the same outcome as that previously reported by Cermak and Levine (1971) for both the Initial Only and Initial plus Post-test Groups (Table I). Within each of these groups the difference in recall following a 5-s presentation was significantly ($P < 0.01$) greater than that following a 2-s presentation. However, there were no significant differences between the two expectancy groups for either presentation rate. This meant that initial recall probability was determined by the length of the presentation interval but not by the knowledge that a post-test recall task would be required. There was a significant difference across trials, the usual result in this type of task, and the probability of correct recall on Trial 6 was significantly below that of Trial 1 ($P < 0.01$).

TABLE I

Per cent correctly recalled across all 6 trials for the 2-s and 5-s presentation interval groups with differential recall expectations

Recall expectations	Presentation interval	
	2 s	5 s
Initial only	57·8	72·8
Initial plus Post-test	55·0	70·0

Post-test Recall

Table II shows the percentage of total words that were correctly recalled during the post-test recall task which followed the completion of all six trials. A significant difference in percentage correctly recalled again existed between the 2-s and 5-s rates ($F = 11·60$, $df = 1,174$, $P < 0·01$.) Inspection of the differences between the 2-s rate and the 5-s rate revealed that the Initial Only and Initial plus Post-test Groups each recalled more words at the 5-s rate than at the 2-s rate ($P < 0·05$), but no such difference existed for the Post-test Only Group. There were no significant differences between expectancy groups when comparisons were made solely within the 2-s presentation rate or within the 5-s rate.

TABLE II

Per cent correctly recalled on the post-test recall task for the 2-s and 5-s presentation interval groups with differential recall expectations

Recall expectations	Presentation interval	
	2 s	5 s
Initial only	32·0	38·5
Initial plus Post-test	32·4	42·4
Post-test only	36·5	37·6

There was a significant difference across presentation trials on the post-test recall with respect to the trial on which a correctly recalled item was initially presented ($F = 19·48$, $df = 5,870$, $P < 0·01$). The Trials × Expectations interaction was also significant, ($F = 5·16$, $df = 10,870$, $P < 0·01$), as was the Trials × Expectations × Rate interaction ($F = 1·80$, $df = 10,870$, $P < 0·05$). Figure 1 depicts the Trials by Expectation interaction for the 2-s rate and Figure 2 shows the same interaction for the 5-s rate. As can be seen the results for the Initial and Post-test Only Groups are essentially comparable for both rates with the Initial plus Post-test results differing depending upon the rate. What is striking about the form of these curves, and what probably led to the Trials by Expectation interaction, is the difference between early trial (Trials 1–3) and later trial (Trials 4–6) recall. Purely for the sake of demonstration the post-test recall results were separated into those correct responses that emanated from presentation Trials 1–3 and those from 4–6. Within the 2-s presentation groups fewer words were
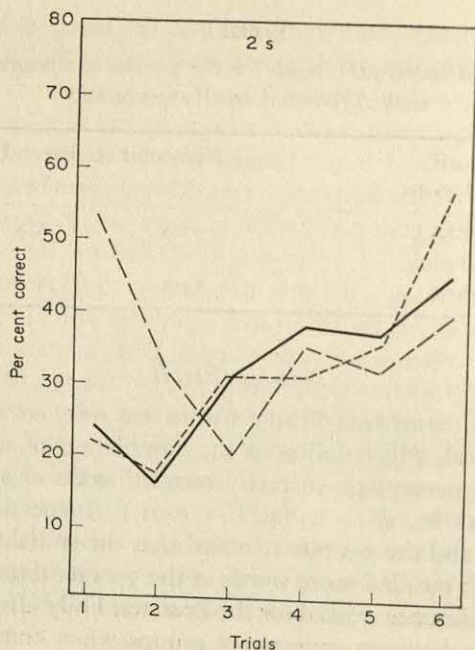
FIGURE 1. Percentage of correct responses on the post-test recall task as a function of the serial position in which the item was presented for the 2-s presentation groups. —— Initial; – – – – Initial-Post; — — — Post.
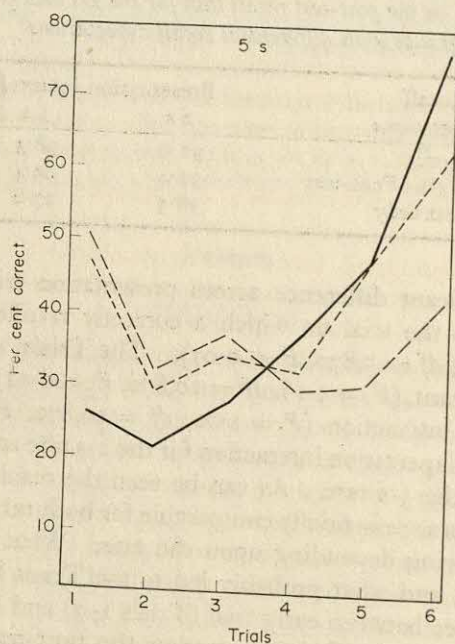
FIGURE 2. Percentage of correct responses on the post-test recall task as a function of the serial position in which the item was presented for the 5-s presentation groups. —— Initial; – – – – Initial-Post; — — — Post.

recalled from the first three trials than from the last three for the Initial Only and Initial plus Post-test Groups ($P < 0.01$) while no such "split-half" difference occurred for the Post-test Only Group. For the 5-s presentation groups the Initial Only Group recalled fewer words ($P < 0.01$) from the first half than the latter half of the list, but the other two groups revealed no split-half difference. Note that as expected only the Initial plus Post-test Group varied from one rate to the other.

The same trend is demonstrated by the order in which the subjects recalled the items on the post-test recall. During post-test recall the items presented on Trial 6 were more frequently among the first three items to be recalled than were items from any other trial position for the Initial Only and Initial plus Post-test Groups ($P < 0.05$). On the other hand, items from Trial 1 dominated the first three items recalled for the Post-test Only Groups ($P < 0.01$). Within the 2-s Initial plus Post-test Group, items from Trial 1 were not recalled any more frequently among the first three recalled items than were any other items, but were recalled more than any other items, except those from Trial 6, for the 5-s Group ($P < 0.05$).

An examination of the words that were correctly given on the post-test revealed that 93% had also been given correctly during the initial recall (all groups combined). The remaining 7% represented words originally omitted during the initial recall but included in the post-test recall task.

Discussion

The performance of the group that did not expect the post-test recall task (the Initial Only Group) is critical in determining which of the three encoding assumptions presented in the introduction is correct. These assumptions made differential predictions with respect to the original presentation positions of the items correctly recalled during post-test recall. The first assumption, which was based on the hypothesis that encoding efficiency declines across trials, predicted a primacy effect. The second assumption, based on the hypothesis that encoding efficiency is constant across trials and that once a trace is formed it does not decay, predicted that neither a primacy nor a recency effect would occur. The third assumption, based on the hypothesis that encoding efficiency is constant across trials but after a trace is formed it decays over time, predicted a recency effect. The results for the Initial Only Group provided the clearest test of these assumptions because the subjects were not anticipating the second recall task. The results on the post-test recall task for this group showed more items recalled from the second half of the list than from the first, and the first items to be recalled came from the final presentation position. Both these results favour the third assumption about encoding, since both indicate that trace strength is a function of the amount of time since the item was originally presented.

The increasing probability of recall across trials for the Initial Only Group on the post-test recall is directly the inverse of the usual trial-by-trial recall probability. Apparently the probability that a particular item will be recalled is a joint function

of the number of similarly encoded traces in memory and the strength of these interfering traces. During trial-by-trial recall the number of similarly encoded traces increases, thus recall probability decreases. However, if time is allowed to pass, the strength of the interfering traces decays and retrieval for the most recently presented items increases. This would explain the primacy effect found in the post-test recall of the present experiment as well as the increasing probability of recall found when the intertrial interval is increased (Loess and Waugh, 1967; Cermak, 1970). Additionally, this experiment has answered the question of whether an item's trace decays with time, or whether it is the interference an item generates that decays with time, in favour of the actual trace decay interpretation.

A striking result occurred when the subjects were told that the second recall task was going to occur. It had been expected that these instructions would produce an increase in the number of items recalled on the post-test recall task. However, no such increase occurred and, in fact, the number of correctly recalled items was the same for all groups regardless of their expectations. What did occur was a shifting in the serial positions in which correctly recalled items were originally presented. The subjects who were asked to recall only during post-test recall displayed a serial-position recall curve that looked similar to a classic serial-position curve. This suggests that the subjects rehearsed the items that had been presented at the beginning of the experiment throughout the rest of the experiment. The group that expected both recall tasks produced different types of curves depending upon the length of the presentation interval. Possibly these subjects had to divide their rehearsal time between the present trial material and prior trial material. Since the only time they had to rehearse was during the presentation interval and/or following their recall and the next item's presentation, the 2-s group was at a disadvantage for two reasons: first, they had less time during the presentation interval; and second, as Cermak and Levine (1971) have shown, less presentation time makes it more difficult to recall the items during the recall interval, thus less time is available for rehearsal of prior items. Consequently the 2-s Initial plus Post-test Group showed a primacy effect that was similar to the Initial Only Group, while the 5-s showed nearly the same serial-position curve as was found for the Post-test Only Group. The recency effect of this 5-s group was somewhat greater than the Post-test Only Group probably because these items had just been recalled on the preceding trial.

In conclusion, it can now be said that encoding efficiency within the context of the distractor technique does not deteriorate across trials. Furthermore, the probability that an item can be retrieved at any time during the procedure is a function of the length of time, or perhaps the number of trials, since it was originally presented. This decay process can be deterred by the subject if he expects to be asked to retrieve the traces a second time. This experiment does raise one interesting question with respect to the trial-by-trial recall scores obtained in this study. These data were not presented because the form of the recall curves by position were identical and have been frequently reported in other studies (e.g. Wickens, 1970). The question that must still be answered is why, if the strength of a trace can be maintained through rehearsal, does this trace not exert more interference than an unrehearsed trace during the retrieval of subsequent items?

References

ADAMS, J. A. (1967). *Human Memory*. New York: McGraw–Hill.

BADDELEY, A. D. and DALE, H. C. A. (1966). The effect of semantic similarity on retroactive interference in long- and short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 5, 417–20.

CERMAK, L. S. (1970). Decay of interference as a function of the intertrial interval in short-term memory. *Journal of Experimental Psychology*, 84, 499–501.

CERMAK, L. S. and LEVINE, R. (1971). Encoding as a function of the presentation-rehearsal interval in short-term memory. *Psychonomic Science*, 23, 423–4.

LOESS, M. and WAUGH, N. C. (1967). Short-term memory and intertrial interval. *Journal of Verbal Learning and Verbal Behavior*, 4, 455–60.

PETERSON, L. R. and GENTILE, A. (1965). Proactive interference as a function of time between tests. *Journal of Experimental Psychology*, 70, 473–8.

PETERSON, L. R. and PETERSON, M. J. (1959). Short-term retention of individual items. *Journal of Experimental Psychology*, 58, 193–8.

WICKENS, D. D. (1970). Encoding categories of words. *Psychological Review*, 77, 1–15.

FORGETTING AS A FUNCTION OF SLEEP AT DIFFERENT TIMES OF DAY

G. R. J. HOCKEY, S. DAVIES AND M. M. GRAY

Department of Psychology, University of Durham

The experiment studied the separate effects of sleep and time period of retention interval on forgetting. A free recall task was given to independent groups of subjects either at night or in the morning, and a second recall demanded 5 h later, after an intervening period of sleeping or waking activity. Oral body temperatures (BT) were measured at each session. The data were analysed in terms of (a) immediate recall at test 1, and (b) amount forgotten from test 1 to test 2. Immediate recall was higher for morning groups, in agreement with previous findings, serial position analysis indicating that the effect is confined to enhancement of the primary memory component. Long-term retention was higher over the night interval, irrespective of sleeping conditions, though having slept at night produced better retention than having stayed awake. Sleep during the morning was not effective in reducing forgetting. BT showed a marked drop for both night groups and rise for day groups over the retention interval. Alternative explanations for the classical sleep/memory findings are suggested in terms of (a) differential effects of sleep stages on memory, and (b) the underlying diurnal variation in BT and other processes.

Introduction

It has often been demonstrated that memory is better after sleep than after a period of normal waking activity (Jenkins and Dallenbach, 1924; Van Ormer, 1932; Ekstrand, 1967; Lovatt and Warr, 1968). Such findings have traditionally been regarded as strong support for the interference theory of forgetting (e.g. McGeoch and Irion, 1952), since sleep can be viewed as effectively removing the subject from the effects of retroactive inhibition (RI) normally experienced in the waking state. Such results can only be interpreted in this way, however, if it can be demonstrated that the lower forgetting rate during sleep is solely the result of this decrease in the amount of interpolated activity.

An important uncontrolled variable in all these studies is the time of day at which sleep occurs. It is now well known that the efficiency of human performance follows approximately a circadian rhythm of the same period and phase as that of many other bodily processes (Conroy and Mills, 1970). Body temperature (BT) in particular has been found to bear a close relation to performance (Kleitman, 1963; Blake, 1967; Hockey and Colquhoun, 1972), both reaching a peak in mid-evening, and attaining their lowest levels during the early hours of the morning. Normally, of course, sleep and BT are highly synchronized; sleep occurs at night when BT is dropping to a low level, and waking activity during the daylight hours

when it is rising. The effects of sleep and of the underlying level of bodily activity are thus usually confounded. BT will be regarded as representative of the general state of physiological activity for the purposes of this paper since it has a stable 24-h periodicity, and has the advantage of being comparatively easy to measure. It should be remembered, however, that while the BT rhythm is fairly typical, the exact phase and shape of the rhythm may be somewhat different in other relevant physiological processes.

All previous studies have compared forgetting over periods of nocturnal sleep and daytime waking. The experimental situation is clearly far too complicated for such a simple design to be of much value. In addition to the sleeping/waking variable, there are differences between the two conditions in the amount of overt activity, the general level of physiological activity, the times of day at which the testing occurs, and other less obvious differences, any or all of which may contribute towards the amount of forgetting. The difficulty is that the variables involved are in the main confounded. As we have said, sleep normally occurs over a night-time period when overt and physiological activity are much reduced. From a theoretical point of view, it is nevertheless important to know what aspects of this complex situation are involved in differential forgetting. One way of approaching the problem would be to arrange for subjects to sleep during the daytime rather than at night; this would at least remove differences in time of testing. In addition, studies of experimentally induced changes in sleep/waking patterns (Colquhoun, Blake and Edwards, 1968; Chiles, Alluisi and Adams, 1968) indicate that physiological rhythms, including temperature, do not normally adapt to the altered routine until several days. This "dissociation" means that in normally nocturnal sleepers the general level of BT and other processes should not be markedly different in subjects who are awake or asleep over a single daytime period. Overt bodily activity and level of EEG arousal are still confounded, of course, but such a situation will at least enable a comparison to be made between effects of sleep *per se* and the time period over which retention occurs. In addition to the usual conditions of daytime waking and night sleep, the present design thus includes those of daytime sleep and night waking.

Several previous experiments have found a relationship between memory and time of day (Blake, 1967; Baddeley, Hatter, Scott and Snashall, 1970). Although the majority of performance functions tend to follow the temperature rhythm, improving steadily through the day, the evidence from these studies suggests that memory is exceptional, being better early in the day when BT is low, then falling off. Since BT and many other rhythms are lowest at night we might therefore expect memory to be most efficient during this period. The studies mentioned above used tests of immediate memory, whereas our primary interest is in retention over long periods. As far as we know, the effect of time of day on long-term retention has not been studied. It would therefore seem worthwhile to examine both these aspects of retention in the same experiment.

The aim of the present study, then, is to investigate some aspects of the relationship between diurnal rhythm and sleep on the one hand, and retention processes on the other. A useful task for this would seem to be free recall, since it gives a characteristic serial position (SP) curve which has been shown to represent the

combination of two components (Glanzer and Cunitz, 1966).   The last few items in the list (recency) can be regarded as being processed through a separate mechanism (primary memory) from earlier items.   These are more stable, and can be thought of as being in long-term (secondary) memory.   From previous work it is known that immediate recall is better in the morning than at night.   The SP data may be analysed for differences between night and morning in either of these two components.

Our primary interest is in forgetting over a long period of time.   Free recall is appropriate for this since it ensures equal exposure time for different groups. Lovatt and Warr (1968) noted that for the more usual paired-associate criterion procedure, subjects tend to need more trials to criterion in the evening (before sleep) than in the morning.   This additional practice could itself partly account for the superiority of sleep in these studies.   One possible disadvantage of free recall, of course, is that initial recall *performance* may be unequal for different groups. This problem may be overcome by estimating relative amounts of forgetting from test to retest in terms of a percentage loss measure.

Method

Subjects and design

The subjects were 40 female undergraduates at Durham University.   All had quite normal sleeping patterns (i.e. neither habitually staying awake all night, nor sleeping all morning).   They were allocated to the four treatment groups: Night Sleep (NS), Night Waking (NW), Morning Sleep (MS) and Morning Waking (MW), 10 subjects to each condition.   The design was thus a 2 × 2 factorial, with independent groups.

Procedure

Subjects were tested in small sub-groups of two or three both before and after a period of 5 h sleeping or waking activity.   They were presented with a list of 30 common bisyllabic nouns, recorded on a tape recorder at a rate of 1 word/2 s, and told to write down as many as possible on completion of the list.   The words could be written in any order with no time limit.   The presentation order of the words was varied randomly for each sub-group, and balanced as far as possible across conditions.   Figure 1 shows the sleep/waking and testing schedules for each of the conditions.   The initial test took place either at 2300 or at 0630 h, and the second test at 0400 or 1130 h, for night groups and day groups, respectively.

After the first test NS subjects went to bed, while NW subjects stayed awake reading, talking or playing games.   Subjects in the morning groups had to be awakened for the initial test from their normal sleep.   In order to minimize the possible effects of sudden awakening (Wilkinson and Stretton, 1971) they were given 5 min to become fully awake, and, in common with all groups, allowed a hot drink (tea or coffee) before the test began.†   MS subjects returned to bed after the test, while MW subjects stayed up, eventually going about their usual morning activities (reading, lectures, shopping, etc.).   All were informed that they would be given another test in 5 h, but were led to believe that it would be of a different nature to the first (this was in order to prevent active rehearsal, as far as possible).   Those in the sleep conditions were instructed to sleep as much possible, and not to read or listen to the radio.   As they had been wakened so early, they found little difficulty in going back

†The design is, of course, only a partial design, since it omits control groups (a) sleeping before the two night conditions, and (b) awake before the two daytime conditions.   These additional conditions could not be included because of the scarcity of subjects.   It was hoped that the consequent loss of pre-test control would be minimized by this standard orienting procedure for all groups at each of the tests.
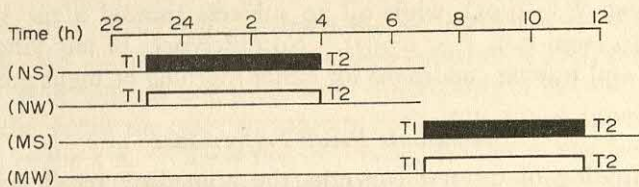
FIGURE 1. Sleep/waking and testing schedules for each group: see text for further details.

to sleep. After a 5-h interval subjects were asked to write down as many of the original words as possible. BT was measured sub-lingually at each session, before any hot drinks, and at least 15 min after any drink, principally to check on the actual trend in the rhythm over the retention interval. The two time periods chosen were expected to coincide roughly with the large downward and upward trends in BT normally found during the night and morning, respectively (Colquhoun *et al.*, 1968).

Results

The main findings are shown in Table I. This gives mean recall performance at each session, and the loss (both absolute and percentage) over the retention period, as well as mean BT levels and changes in BT over the 5-h interval.

TABLE I

Measures of recall performance and body temperature in test 1 and test 2

Group	Recall		Absolute loss	Percentage loss*	Body temperature		(° C) Change
	Test 1	Test 2			Test 1	Test 2	
NS	8·9	6·8	2·1	22·8	36·8	36·0	−0·8
NW	9·4	6·3	3·1	34·5	36·6	36·0	−0·6
MS	10·4	5·8	4·6	43·0	36·2	36·8	+0·6
MW	11·2	5·6	5·6	46·9	36·3	36·9	+0·6

*These figures are calculated from individual per cent loss scores, and not from the means shown in previous columns.

The main analysis concerns the amount of forgetting from the first to the second sessions. This can be taken either as the absolute or percentage loss. The latter is more appropriate in this case since the initial recall scores are not equal, and is in any case a more conservative measure of differences between the treatments. An analysis of variance carried out on the per cent loss data showed a large main effect of time of day ($F = 17 \cdot 7$, $df = 1$, 36, $P < 0 \cdot 001$) and a non-significant effect of sleeping conditions ($F = 4 \cdot 0$, $df = 1$, 36, $P > 0 \cdot 05$). The interaction was non-significant ($F = 1 \cdot 0$, $df = 1$, 36, $P > 0 \cdot 05$). Differences between the treatments were further analysed using the Newman-Keuls procedure (Winer, 1962). This revealed significant differences at the 0·01 level between NS and MS and between NS and MW. No other comparisons were significant.

The temperature data were clear-cut, BT being higher for the combined night groups at test 1 ($t = 2 \cdot 84$, $df = 39$, $P < 0 \cdot 01$) and lower at test 2 ($t = 4 \cdot 16$, $df = 39$, $P < 0 \cdot 01$) than for the combined morning groups. In the night groups 15 out of 19 subjects showed a drop in BT (1 showed no change) over the 5-h

period (sign test, $P = 0.01$), while all 20 subjects showed a rise in BT in the morning groups (sign test, $P < 0.001$). No differences of any kind were found between sleep and waking conditions for either morning or night groups.

Immediate Recall Performance

The other finding of interest concerns the immediate recall performance of morning and night groups. The mean recall scores for test 1 are higher for combined morning groups than for night groups ($t = 2.67$, $df = 39$, $P < 0.02$). This is in agreement with the superior immediate memory performance associated with low points in the diurnal rhythm, mentioned earlier. Figure 2 shows the SP curves for each of the two combined groups (both $N = 20$). The only apparent discrepancy between them is in the last few positions, where the morning curve is generally higher (in six of the last seven points). This suggests that the effect is
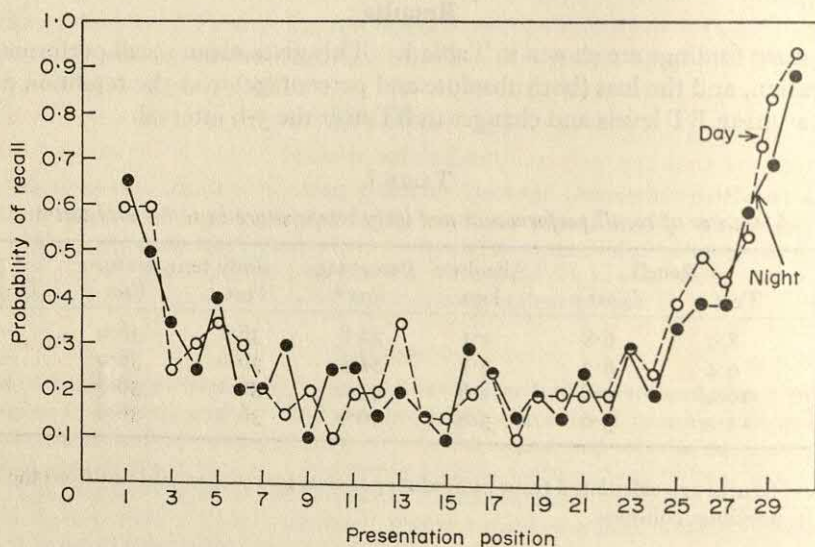
FIGURE 2. Serial position curves for combined night groups and morning groups on Test 1.

restricted to an enhancement of primary memory, rather than representing an overall improvement in immediate recall efficiency.

Discussion

Within the limitations of the experimental design the present results clearly cannot be interpreted in terms of any simple "low-interference" effect of sleep. Only when sleep occurs at night is there the usual substantial reduction in forgetting. The typical comparison in this kind of research has been between conditions that approximate our NS and MW groups, and these do indeed show the largest difference in forgetting. Failure to control for the time at which sleep occurs has resulted, however, in an oversimplification of this finding. Sleep during the day, though it fulfils the condition of removing the subject from sources of interference,

makes no appreciable difference to the amount of forgetting that takes place. It could, of course, be argued that subjects in the MS condition did not sleep as much as those in the NS group. Though no strong evidence is available on this point, checks made by invigilators and subsequent questioning of the subjects lead us to believe that this is not the case: all subjects in the MS group appear to have slept for at least 4 of the 5 h. This is not surprising, in view of the fact that they had been wakened about 2 h before their normal rising time, so were still tired.

The data are consistent with two alternative kinds of explanation, though these may not be mutually exclusive. The experiment was designed specifically to compare the effects of sleeping or waking with those of time of day, and from the analysis of variance it is clear that time of day is the more important factor. Retention over the early part of the night is superior to that over the morning period, whether subjects are sleeping or not. The BT data confirm that these two periods are associated with pronounced downward and upward swings, respectively, in the diurnal cycle. The difference between NS and NW may be due to additional stress brought about by sleep deprivation, though retention for NW subjects is, nevertheless, somewhat higher than for those in either of the morning groups.

There remains the possibility, however, that while daytime sleep is not beneficial to memory, sleep at night does have a genuine effect. One way in which this may occur is that the preservation of stored information is dependent upon the qualitative nature of sleep. During the early part of the night (as in our NS condition) the EEG record shows predominantly large-voltage, low-frequency waves (known as stage 4 or slow-wave sleep), while later in the nocturnal sleep period and during the early part of the day (when subjects are made to "oversleep") the dominant characteristic is low-voltage, high-frequency activity of the REM type (Webb and Agnew, 1967; Hartmann, 1967). The data may be accounted for, at least in part, if slow-wave sleep was in some way beneficial to memory, while REM sleep was not, since our two sleep periods would be expected to differ in the relative amounts of each kind of sleep. This hypothesis has been tested by Yaroush, Sullivan and Ekstrand (1971). They found reduced forgetting only after sleep for the first half of the night: sleep for the second half was little better than a period of morning waking activity. They interpreted their findings as supporting a theory of slow-wave sleep facilitating the consolidation process, though the two sleep periods again differ, of course, not only in kind of sleep but in the time at which they occur. The effects of the two kinds of sleep on memory can be compared directly by using selective deprivation procedures over the same retention period. Using this approach Ekstrand, Sullivan, Parker and West (1971) failed to obtain any differences between slow-wave and REM deprivation (which produce predominantly REM and slow-wave sleep, respectively), while Empson and Clarke (1970) did find a difference, but in the opposite direction. REM-deprived subjects recalled less than yoked controls, despite having taken more stage 4 sleep.

The evidence for differential effects of sleep stages on memory is thus inconsistent, probably largely due to the difficulty of adequate experimental control, though it is clearly an important area for further study. The present data do not offer a simple alternative explanation in terms of diurnal variation, though it seems

clear that retention is better over the night period, whether subjects are asleep or awake. The BT data (Table I) show evidence of only a slight difference in *average* level between night and morning groups (about 0·2° C). This seems hardly able to account for the considerably better retention in the night conditions. The analysis of immediate recall shows an advantage for morning presentation, when BT is significantly lower, the effect being apparently confined to the last few serial positions. This result confirms those of previous workers (Blake, 1967; Baddeley *et al.*, 1970), and implies a change in the capacity of primary memory at different times of day.

One puzzling feature of the BT data is that although the high BT at night leads to poor immediate recall it does not result in a more rapid memory loss following the recall test and a subsequently lower level of retention: this might be expected since most forgetting is known to occur in the period immediately after presentation. It may be that storage is somehow facilitated by the fact that BT is *dropping* over this period. In the morning groups it starts off lower and *rises* by an almost equal amount. A second possibility is that the higher BT at night results in better consolidation of stored material. There is no basis for distinguishing between these two explanations in the present study, and it is clearly possible that the observed difference is more directly due to diurnal variation in some other bodily process. While we have used BT as an index of the level of physiological activity, it must be remembered that other processes may exhibit rhythms of a rather different phase. We do not yet know which of these are involved in storage and decay. It would nevertheless be useful to compare forgetting over periods when BT is either rising or falling, but starting at the same level, or to compare phases of the BT cycle having different initial levels but approximately equal changes.

As we suggested earlier, the situation is one of considerable complexity, with many interacting and confounding variables. All that we have been able to do in this experiment is to demonstrate that sleep *per se* does not seem to be crucial to the reduced forgetting found at night, and is of no benefit at all when taken in the daytime. The precise relationship between sleep patterns, diurnal rhythms and different components of the memory process have hardly been touched on as yet, and clearly require much further experimental analysis. One obvious necessity is to check that these findings hold for the more traditional paired-associate paradigm, since free recall may be rather a specific kind of memory task. On a broader front there is plainly much important information to be gained from experimental shift-work and trans-zonal displacement studies: what happens to memory as sleep patterns and other rhythms adapt to the changed routine? So far, this aspect of performance has received little attention in this field.

References

BADDELEY, A. D., HATTER, J. E., SCOTT, D. and SNASHALL, A. (1970). Memory and time of day. *Quarterly Journal of Experimental Psychology*, **22**, 605–9.

BLAKE, M. J. F. (1967). Time of day effects on performance in a range of tasks. *Psychonomic Science*, **9**, 349–50.

CHILES, W. D., ALLUISI, E. A. and ADAMS, O. (1968). Work schedules and performance during confinement. *Human Factors*, **10**, 143–96.

COLQUHOUN, W. P., BLAKE, M. J. F. and EDWARDS, R. S. (1968). Experimental studies of shift work, I: a comparison of "rotating" and "stabilised" 4-hour shift systems. *Ergonomics*, **11**, 437–53.

CONROY, R. T. W. L. and MILLS, J. N. (1970). *Human Circadian Rhythms*. London: Churchill.

EKSTRAND, B. R. (1967). Effect of sleep on memory. *Journal of Experimental Psychology*, **75**, 64–72.

EKSTRAND, B. R., SULLIVAN, M. J., PARKER, D. F. and WEST, J. N. (1971). Spontaneous recovery and sleep. *Journal of Experimental Psychology*, **88**, 142–4.

EMPSON, J. A. C. and CLARKE, P. R. F. (1970). Rapid eye movements and remembering. *Nature*, **227**, 287–8.

GLANZER, M. and CUNITZ, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior*, **5**, 351–60.

HARTMANN, E. L. (1967). *The Biology of Dreaming*. Springfield, Illinois: Charles C. Thomas.

HOCKEY, G. R. J. and COLQUHOUN, W. P. (1972). Diurnal variation in human performance: a review. In COLQUHOUN, W. P. (Ed.), *Aspects of Human Efficiency: Diurnal Rhythm and Loss of Sleep*. London: English Universities Press.

JENKINS, J. C. and DALLENBACH, K. M. (1924). Oblivescence during sleep and waking. *American Journal of Psychology*, **35**, 605–12.

KLEITMAN, N. (1939). *Sleep and Wakefulness*. New York: University of Chicago Press.

LOVATT, D. J. and WARR, P. B. (1968). Recall after sleep. *American Journal of Psychology*, **81**, 253–7.

McGEOCH, J. A. and IRION, A. L. (1952). *The Psychology of Human Learning*. New York: Longman, Green & Co.

VAN ORMER, E. B. (1932). Retention after intervals of sleep and waking. *Archives of Psychology*, **21**, 1–49.

WEBB, W. B. and AGNEW, H. W. (1967). Sleep cycling within 24-hour periods. *Journal of Experimental Psychology*, **74**, 158–60.

WILKINSON, R. T. and STRETTON, M. (1971). Performance after awakening at different times of night. *Psychonomic Science*, **23**, 283–5.

WINER, B. J. (1962). *Statistical Principles in Experimental Design*. New York: McGraw-Hill.

YAROUSH, R., SULLIVAN, M. J. and EKSTRAND, B. R. (1971). Effect of sleep on memory. II: Differential effect of the first and second half of the night. *Journal of Experimental Psychology*, **88**, 361–6.

SHORT-TERM VISUAL MEMORY AND PATTERN MASKING

D. C. MITCHELL†

Department of Psychology, University College London

The results of two experiments involving the matching of unfamiliar, nameless shapes (Gibson forms) indicated that a visual representation of a brief (30–100 ms) stimulus survives in a Short-Term Visual Memory (STVM) for 6 s or more after the onset of a pattern mask. On the basis of these results a possible alternative to Sperling's (1967) model for short-term memory for visual stimuli was presented. In this model it is assumed that recognition processes occupy several hundred milliseconds and continue after the arrival of the mask using the information available in STVM.

Introduction

Over the last decade a variety of experimental approaches has provided support for the notion of high-capacity visual storage lasting less than a second (Sperling, 1960; Averbach and Coriell, 1961; Eriksen and Collins, 1967, 1968; Haber and Standing, 1969, 1970). This brief form of visual persistence has been called the icon by Neisser (1967).

In an extension of his earlier work, Sperling (1963, 1967) presented subjects with arrays of letters followed after short intervals by a pattern mask of random letter fragments. He found that the number of letters reported rose sharply with the stimulus duration until a limit of about four or five after which performance levelled off. According to the model which he put forward to account for these data the stimulus information is first registered in iconic storage and then rapidly read into a symbolic buffer store, the Recognition Buffer, where it is held pending conversion into an acoustic form for maintenance by rehearsal in an Auditory Information Store (AIS). Sperling argued that the effect of the mask was to erase the icon and therefore to terminate the process of reading out from the icon to the Recognition Buffer. Thus performance in his tachistoscopic recognition task improved with stimulus duration because more items could be identified and rescued from the icon before the erasure. The asymptote in performance was attributed to the capacity limitations of AIS.

Assuming serial processing of the items Sperling (1963) estimated that the read-out rate from the icon was 100 items/s. For parallel processing a corresponding estimate would be at least 20–25 items/s since a single item can be correctly identified when the stimulus duration is 40–50 ms (Liss, 1968). According to the model

†Present address: Department of Psychology, University of Exeter

this process entails stimulus identification since the contents of the Recognition Buffer are symbolic and non-visual.

The validity of these estimates of recognition rate depends on the assumption that the icon is erased within a few milliseconds of the arrival of the mask. This assumption has been supported by the results of experiments in which subjects set the time of an auditory click to coincide with the subjective disappearance of the image (Sperling, 1967; Haber and Standing, 1970). For stimulus duration of up to 100 ms the click settings were typically within about 20 ms of the onset of the mask. However, the click experiment has been criticized on the ground that subjects may synchronize the click with the onset of the mask while visual information survives well beyond this point (Neisser, 1967; Kahneman, 1968).

Data which support this suggestion have been presented by Posner, Boies, Eichelman and Taylor (1969). They argued that if a match between two letters of the same shape (*physical match*) is faster than one in which they only have the same name (*name match*) then a visual description (Visual Code) of the first letter must have survived until the presentation of the second. They found that such a difference occurred when the two letters were separated by up to 2 s even if the first was followed by a mask consisting of visual noise. With the 2-s interval there was no time difference, presumably because the *name match* always occurred first, and so no further delay of the Visual Code could be detected. However, Phillips and Baddeley (1971) showed that the persistence continues for more than three seconds after a mask. This conclusion was based on the accuracy with which subjects were able to match two 5×5 matrices of black and white squares with an intervening random matrix as a mask. A similar conclusion was reached by Taylor (1969), who found that the matching latencies for visual stimuli separated by a 3-s mask were unaffected by the namability of the stimulus materials.

If subjects are able to use this post-mask visual information then Sperling's estimate of recognition rate may not be valid. However, some caution is needed before drawing this conclusion since it is not clear that the visual information would be available under the conditions of the tachistoscopic recognition task where, for example, the stimulus duration was considerably shorter than the 0·5 s or more employed above. This generalization might have been justifiable if the Visual Code were the same as the icon; however, the evidence suggests that it is not. Its duration is considerably longer than estimates of iconic persistence, and it is impaired by a distracting task (Posner *et al.*, 1969) whereas the icon is not (Doost and Turvey, 1971). It may be more appropriate to identify the Visual Code with the Short-Term Visual Memory (STVM) surviving 25 s or more which has been studied by Kroll, Parks, Parkinson, Bieber and Johnson (1970) and Warrington and Shallice (1972). If this is the case visual information may survive beyond the mask only if there is sufficient time during the stimulus duration for it to be registered in STVM.

It is unlikely that the methods available could be used to determine whether this is possible with short stimulus durations. The complexity and unconnectedness of 5×5 matrices may be too great to allow matching performance above chance and in Posner's task the reduction in stimulus duration may increase the time for the *physical match* more than that for the *name match*, destroying the basis

for inferences about the *Visual Code*. A further difficulty when letters are used as stimuli is that the visual information surviving the mask may be wholly or partially retained by means of the activation of a pre-existing visual prototype, template or abstract representation of the stimulus. Since this form of storage would itself involve stimulus categorization it would be inappropriate to use such evidence in an argument against rapid recognition.

The experiments in this study were designed to determine whether visual information persists after the onset of a pattern mask when the stimulus duration is very short and when there is no visual prototype of the stimulus form to allow the type of storage outlined above. The technique employed involved the matching of unfamiliar, but letter-like nameless shapes (Gibson forms) which were presented sequentially.

Experiment I

Apparatus

The apparatus used was a semi-automatic four-field tachistoscope. Two of the fields contained spools of 35 mm film on which the stimuli were photographed. The film was transported through the gates, frame by frame, by a Mullard synchronous motor. The remaining two fields each contained a single 35 mm slide. The distance from the eye to the gate was 72·4 cm and the image was magnified $2\frac{1}{2}$ times by a pair of lenses at the eyes.

Lighting was by back-projection using a zig-zag mercury vapour fluorescent lamp (rise time 300 μs, fall time 100 μs) behind a 2 mm plate of ICI Perspex Opal 042. The light spectrum of this source peaked at wavelengths of 425–450 nm.

Each lamp was fired by an independent Driver Module, a high voltage D.C. source, to provide the necessary lamp-triggering voltage (850 V) and running voltage (380 V). Each Driver Module was under the control of a valve whose state was determined by a high-speed reed relay controlled by the solid state logic timing circuits (see below).

The lamp cycling sequence was controlled by solid state logic. The cycle could be started either by the subject using a hand switch or by the experimenter and was terminated on completion if a "stop" button was pressed during its execution but otherwise repeated itself indefinitely.

The duration of each component of the cycle was controlled by a separate monostable timer (Texas SN 74121N high-performance, temperature-compensated TTL monostable) which could either be set to 40 ns, in which case the lamp-illuminating reed relay had insufficient time to operate, or varied continuously between 6 ms and 30 s. The time was determined by one of three capacitors (for coarse steps) and a variable resistor (for fine increments) across two external terminals of the timer. This arrangement gave the following continuous time ranges: (1) 6–128 ms; (2) 60 ms–1·2 s; (3) 0·5–20 s. An oscilloscope was used to calibrate the timer dials directly in terms of light output pulse durations.

Procedure

At the beginning of each trial the adaptation field was replaced by a brief presentation of a single shape, the Stimulus form. This was followed by a pattern mask which was then replaced after varying intervals by a second shape (the Probe form) together with the same mask. The cycle ended with the return of the adaptation field (see Fig. 1). The subject's task was to say whether the two shapes were same or different.

Two types of material were used; numerals ("two" to "six" inclusive) and three sets of 5 nameless shapes (12 from Gibson (1965) and 3 additional similar forms). All stimuli were constructed from features of a Type 6 Uno Penstencil and inked in black on a white background. They were then photographed in Ilford Microneg 373 35 mm film and developed in Johnson Copyphen to give maximum contrast. A positive copy was used in the tachistoscope.
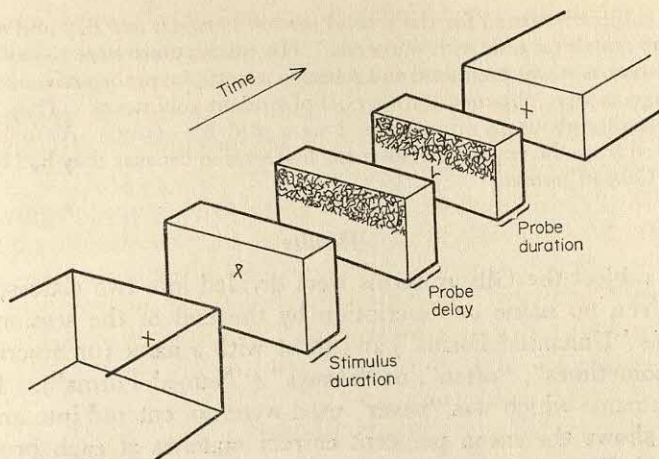
FIGURE 1. The stimulus sequence for the matching task.

The adaptation field was blank except for a focusing cross at the centre subtending an angle of 0·1° at the eye. The Stimulus form, which covered between 0·3° and 0·6° vertically and horizontally depending on the form, was located with its centre 0·6° above the centre of the slide. The pattern mask, consisting of a dense array of numeral and letter fragments (Sperling, 1963), occupied only the top half of the field, leaving the bottom half blank. The Probe slide consisted of a single item equal in size to the Stimulus form appearing 0·6° below the centre together with the pattern mask covering the top half of the field. The stroke width of the focusing cross was approximately 0·02°, and 0·08° for all other lines. The luminance of each blank field was 9·4 cd/m².

The 120 pairs of items to be matched were made up of three blocks of 40 in which each item was paired with each other once and with itself four times. The order in which the pairs were presented was otherwise random with two different randomizations for each Gibson form set.

For each subject the stimulus duration was either 30 ms or 40 ms for the Gibson forms and 20 ms for the numerals. There were three probe delays, 0·5, 3·0 and 6·0 s, and the probe duration was always 1 s. Each probe delay occurred equally often with each stimulus-probe combination, but otherwise their order was random.

On each trial the stimulus cycle was started about ½ s after the sound of the film change mechanism, which was used as a non-verbal "ready" signal. Response latencies were measured using a voice key and a Watesta timer, and feedback was given after blocks of 10 trials.

At the beginning of the session the subject spent a few minutes adjusting the focus and convergence of the eyepieces to obtain a single sharp image of the cross. This was followed by a practice block of 40 trials with Gibson forms to familiarize him with the apparatus and task. He was instructed to "try to hold a visual picture" of the stimulus form as accurately as possible so that he could compare it with the probe form. Immediately after the practice was a 120-trial Experimental block using a different set of Gibson forms. There were short rests after each 40 trials.

At the end of the session the subject was asked whether he remembered each form visually or as a name, and was required to write any name or descriptions he had invented against drawings of the forms on the duplicated response sheet. His attention was then drawn individually to each of the unnamed forms and he was asked whether he was "perfectly sure" that he had made up no names for it. If this had no effect commonly used names were suggested to him to help stimulate the recall of any names which might have been forgotten. The subject then indicated for each form whether he had used a name "never," "sometimes", "often" or "always".

When the subject returned for the second session between one day and a week later he performed the matching task with numerals. His instructions were to say the "number" to himself as soon as it was presented and rehearse it until the probe arrived.

The 24 subjects were obtained from a pool of student volunteers. They participated in two sessions lasting about 45 min each and were paid 10/– (50p). A further six subjects were eliminated from the experiment after the first session because they had invented names for all of the Gibson frames.

Results

For each subject the Gibson forms were divided into two classes, those which had been given no name or description by the end of the session (henceforth referred to as "Unnamed Forms") and those with a name (or description) which was used "sometimes", "often", or "always" ("Named Forms"). Forms which were given a name which was "never" used were not entered into any analysis.

Figure 2 shows the mean per cent correct matches at each probe delay for numerals and Unnamed Forms (both stimulus and probe forms Unnamed). Figure 3 shows the corresponding data for (both) Named and (both) Unnamed
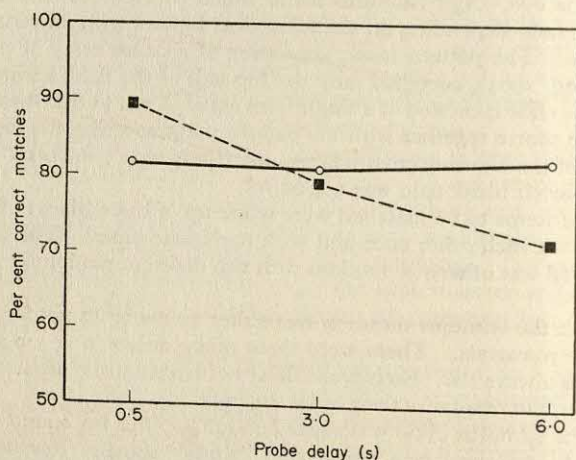
FIGURE 2. Mean per cent correct matches for Named Forms and Unnamed Forms for each stimulus duration and probe delay. ○———○ Numerals; ■ – – – ■ Unnamed Forms.

Forms for each of the two stimulus durations. It should be pointed out that the data plotted for Named Forms may not be strictly comparable with those for Unnamed Forms or numerals since only 18 subjects had Named Forms whereas all 24 subjects contributed to the graph for the numeral condition and to one of the two Unnamed Form graphs.

The data for proportion-correct matches were transformed, using the inverse circular transformation, and entered into a repeated-measures Analysis of Variance (ANOVA) of stimulus duration × Gibson form set × probe delay × name (numerals or Unnamed Forms) × subjects. Performance differed according to the Gibson form set ($F = 4 \cdot 17$, $df = 2$, 18, $P < 0 \cdot 05$) presumably because of variations in form similarity within the sets. The effects of probe delay ($F = 3 \cdot 95$,

$df = 2, 36, P < 0.05$) and name $\times$ probe delay interaction ($F = 3.93$, $df = 2, 36$, $P < 0.05$) were also significant and are examined in more detail in the trend analysis below. The only other significant effect was the interaction of name and stimulus duration ($F = 6.59$, $df = 1, 18$, $P < 0.05$) which arose simply because the stimulus duration was varied for Gibson forms but not for numerals. All other effects failed to reach significance, in particular the F-ratio for the name factor was less than unity, indicating that the choice of stimulus durations for the numeral and Gibson form conditions succeeded in placing the performance curves in the same part of the scale. Consequently, distortions due to extremity effects should be minor.

A trend analysis of the data showed that performance fell with probe delay (the linear component of the probe delay effect was significant ($F = 5.50$, $df = 1, 18$, $P < 0.05$)). There was no evidence of deviations from linearity ($F < 1$, $df = 1, 18$). The rate of fall for Unnamed Forms was greater than that for numerals.
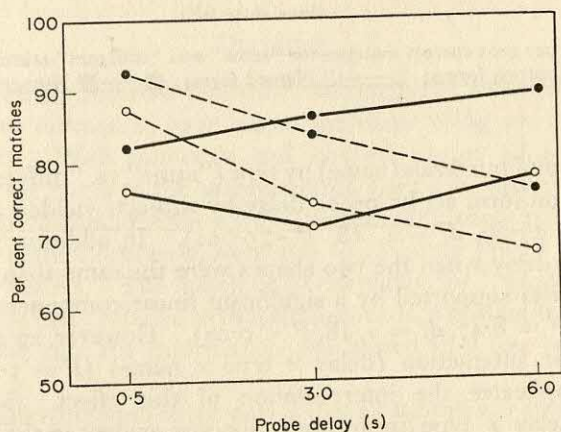
FIGURE 3. Mean per cent correct matches for numerals and Unnamed Forms for each probe delay. — — — Unnamed Forms; ——— Named Forms; ● 40 ms; ○ 30 ms.

This was indicated by the linear component of the probe delay $\times$ name interaction ($F = 7.34$, $df = 1, 18$, $P < 0.025$). Again there was no evidence of curvilinearity ($F < 1$, $df = 1, 18$). The matching performance with numerals showed no evidence of falling with probe delay ($F < 1$, $df = 1, 18$, not independent of the tests above).

A second ANOVA on arcsine-transformed data with Named vs. Unnamed Forms (name) $\times$ probe delay $\times$ subjects was performed using the data of the 18 subjects who had both Named and Unnamed Forms. The linear component of the probe delay $\times$ name interaction was again significant ($F = 7.69$, $df = 1, 17$, $P < 0.025$), indicating that Named and Unnamed Forms were retained differently.

In Figure 4 the performance data are plotted separately for trials on which the stimulus and probe forms were the same and those when they were different. An ANOVA performed on the arcsine-transformed proportions of correct matches for
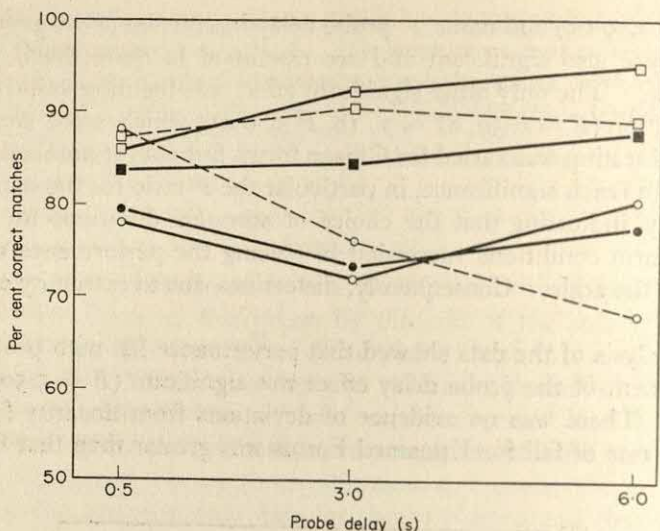
FIGURE 4. Mean per cent correct matches for "same" and "different" stimulus/probe combinations. ○ – – – □ Unnamed forms; ○——□ Named forms; ●——■ Numerals; ■ □ Different; ● ○ Same.

Unnamed Forms and numerals (name) by type ("same" vs. "different") by stimulus duration by Gibson form set by probe delay by subjects yielded a significant type main effect ($F = 17.29$, $df = 1, 18$, $P < 0.001$). In addition performance fell faster with probe delay when the two shapes were the same than when they were different. This was supported by a significant linear component of the delay × type interaction ($F = 8.47$, $df = 1, 18$, $P < 0.01$). However, an almost significant third order linear interaction (delay × type × name) ($F = 3.47$, $df = 1, 18$, $< P < 0.1$) complicates the interpretation of this effect. Separate analyses showed that the delay × type linear interaction was evident in the Unnamed Form
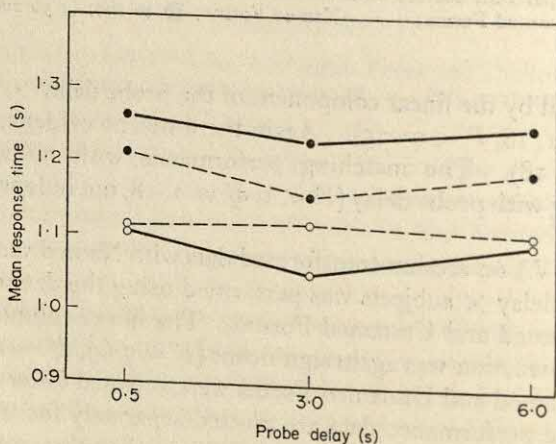
FIGURE 5. Mean correct same and different response times for Unnamed Forms and numerals. – – – Unnamed Forms; —— Numerals; ● Different; ○ Same.

condition, but did not approach significance in the numerals condition ($F = 6 \cdot 90$, $df = 1, 18, P < 0 \cdot 025$; $F = 2 \cdot 60$, $df = 1, 18, P < 0 \cdot 1$, respectively).

At $0 \cdot 5$ s probe delay all 24 subjects performed better than chance at the matching task with Unnamed Forms. The corresponding frequencies for the 3-s and 6-s probe delays were 21 better than chance, 2 worse and one tie; and 19, 3, with two ties. All of these were significant on a binomial test ($P < 0 \cdot 001$).

Figure 5 shows the mean correct "same" and "different" response latencies at each probe delay when the stimuli were Unnamed Forms and numerals.

The mean RTs were transformed (log transformation) and entered into a repeated measures ANOVA for stimulus duration × Gibson form set × probe delay × name × type (same or different) × subjects. The only effect to reach significance was the type main effect ($F = 24 \cdot 7$, $df = 1, 18, P < 0 \cdot 001$) which indicated that "same" responses were faster than "different" responses. The mean RT for numerals ($1 \cdot 173$ s) was clearly no different from that for Unnamed Forms ($1 \cdot 174$ s) ($F < 1$, $df = 1, 18$).

Discussion

Of the 24 subjects 22 reported that they had retained Unnamed Forms visually. The remaining two claimed to have traced the shape using eye and finger movements respectively. With numerals and Named Forms all subjects reported having rehearsed the name of the stimulus until the probe arrived.

The matching performance data were consistent with these reports. As would be expected with perfect rehearsal, numerals and Named Forms showed no fall in performance over the 6 s. The matching performance for Unnamed Forms, however, fell with probe delay relative to that for numerals, a finding which is consistent with the suggestion that the information was stored in a less rehearsable visual form.

An alternative interpretation of the data is that all stimuli were coded and retained as verbal descriptions but for numerals and Named Forms these descriptions were succinct (e.g. one word) while for Unnamed Forms they were longer and therefore less rehearsable. These differences in ease of rehearsal could lead to perfect retention in the former cases but loss of information for "Unnamed Forms". This explanation seems unlikely considering the thoroughness of the screening procedure used to isolate the "Unnamed Forms" and receives no support from the response latency data which failed to reflect the postulated differences in description length for numerals and "Unnamed Forms".

The evidence suggests, therefore, that the Unnamed Forms were retained visually for at least 6 s after the onset of the mask. The fall in matching performance with probe delay may be attributed to a decline in the "amount of visual information" in storage, although the data did not permit any exact specification of the meaning of this term. Linear extrapolation of the decay curve, which may be partially justified by the absence of deviations from linearity in the trend analysis, gave estimates of visual persistence of 11 and 14 s for the stimulus durations of 30 ms and 40 ms respectively. Since the decay curve is more likely to be negatively accelerated it is probable that the true persistence of visual information is longer.

The conditions of the experiment favoured long survival of visual information since the subject was able to devote his attention to the rehearsal of a single shape. Also subjects with good visual memories are less likely to have resorted to naming the Gibson forms and therefore less likely to have been eliminated from the experiment.

Some indication of the nature of the decay process may be obtained from the pattern of results in the Unnamed Form condition. For example, it would be difficult to reconcile the differential fall in performance on the "same" and "different" trials with an all-or-none model of decay. The guessing bias towards "different" on the "none" trials would have to increase with delay to produce such a result. A more feasible alternative is that delay involves the loss of elements from a multicomponent-type feature list. In this case the data can be accounted for if it is assumed that more detailed visual information is required for a successful "same" than for a "different" match.

Experiment II

The previous experiment did not exclude the possibility that the information surviving the mask was retained by means of a visual prototype. Although the forms were initially unfamiliar, repeated presentation may have given the subjects an opportunity to construct prototypes for use in the task.

The current experiment was designed to test whether visual information can survive after a mask when the stimulus form has never previously been seen by the subject. In these circumstances no appropriate visual prototype can be available.

Method

The stimulus sequence and the matching task were the same as in Experiment I. After focusing and adjusting the tachistoscope, each subject performed a practice block of 20 trials in which upper case letters ("A" to "E") were used as stimulus and probe items. The subject was then told that in the experimental block the letters would be replaced by novel shapes but the task would otherwise remain unchanged.

The experimental block consisted of only 10 trials. Five of the Gibson forms were used for five "same" trials and the remaining 10 for five "different" trials. There were six randomizations of stimulus/probe form combinations and trial orders. Each subject was allocated one of 2 stimulus durations (40 or 100 ms) and one of 3 probe delays (100 ms, 0·5 and 6·0 s). The probe duration was fixed at 1 s. Six subjects were run in each combination of the conditions. The luminance of each blank field was 9·4 cd/m².

The 36 subjects were obtained as before and paid 2/– (10p) for a 20 min session.

Results

Figure 6 shows the mean per cent correct matches at each probe delay. One sample t-test on these data (after arcsine transformation) showed that matching performance was significantly better than chance for both stimulus durations ($t = 2·21$, $df = 17$, $P < 0·05$; $t = 13·56$, $df = 17$, $P < 0·001$).

An ANOVA (probe delay × stimulus duration × subjects) yielded a highly significant effect of stimulus duration ($F = 84·5$, $df = 1$, 30, $P < 0·001$) but no delay main effect and no delay × stimulus duration interaction ($F < 1$ in each
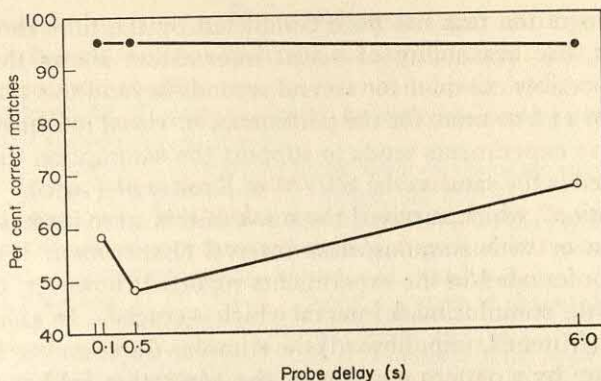
FIGURE 6. Mean per cent correct matches at each probe delay and for each stimulus duration (Experiment II). ●——● 100 ms; ○——○ 40 ms.

case). The absence of delay effects may be attributed in part to a ceiling effect in the 100 ms condition; however, it is also possible that visual information may survive for longer than the data of Experiment I suggested.

Discussion

The high success rate obtained in the 100 ms condition strengthens the conclusion that visual information may be retained after a pattern mask. Alternative explanations would require that a name or description of the form was rapidly invented during the brief interval before the arrival of the mask. A single subject claimed that he used names for three of the 10 forms, but there is no evidence that these occurred before the mask. All other subjects reported that if they "saw" the stimulus form they remembered it visually.

Since subjects were highly successful at the matching task even without having previously seen the shapes it may be concluded that pre-existing visual prototypes are not crucial for the retention of visual information after a mask. However, they may improve matching performance. This is suggested by a comparison of the performance in the 40 ms condition in the current experiment with that for Un-named Forms in Experiment I where such prototypes may have been available. However, this difference may also be accounted for by differences in task familiarity and the sizes of stimulus ensembles used.

The increase in matching performance with stimulus duration suggests that the "amount of visual information" retained after the mask is determined by the stimulus duration or the stimulus/mask interval.

General Discussion

The data from the two experiments presented confirm that it is possible to retain visual information after the onset of a pattern mask even when the stimulus duration is as short as 30–40 ms. This can occur without pre-existing visual prototypes of the shapes (Experiment II). These results show that there is no need to assume that identification of the contents of the stimulus array in the

tachistoscope recognition task has been completed by the time the mask arrives. On the contrary, the availability of visual information shows that recognition processes could possibly continue for several seconds beyond this point.

The estimate of 11 s or more for the persistence of visual information under the conditions of these experiments tends to support the assumption that the store in which it is retained is the same as the STVM of Kroll *et al.* (1970). The "amount of visual information" which survived the mask in this store increased either with stimulus duration or with stimulus/mask interval (Experiment II). These two measures were confounded in the experiments reported; however, other evidence suggests that it is the stimulus/mask interval which is crucial. In a short pilot study with six subjects (Mitchell, unpublished) the stimulus duration was kept fixed and was followed either by a pattern mask or by the adaptation field and it was found that matching performance with Unnamed Forms was significantly worse in the former condition. If the stimulus duration alone had determined the "amount of visual information" post-stimulus events would have had no effect.

The relationship between matching performance and the stimulus/mask interval may be explained if it is assumed that visual information is transferred from the icon to STVM until the icon is erased, or possibly degraded, by the mask and that the "amount of visual information" placed in STVM is dependent on the time available for transfer.

This account of the data suggests a possible alternative to Sperling's (1967) model for the tachistoscopic recognition task. The initial sharp improvement with stimulus duration may be attributed to increases in the "amount of visual information" in STVM and the levelling of performance at 4·5 items to the capacity limitations of this store. (Experiments with deaf subjects show that this may be 4–5 items under some conditions (Ross, 1969)). The information in STVM would then be classified and converted into an acoustic form for maintenance by rehearsal until the report is made.

This modification differs from Sperling's model in that it is assumed the buffer store is visual rather than symbolic; a point which is crucial for inferences about recognition rate and the effects of masking. According to Sperling's model the mask limits the time available for recognizing the contents of the stimulus array, while in the Visual Buffer model its effect is to control the "amount of visual information" or signal-to-noise ratio in STVM. If the second formulation is correct one implication would be that pattern masking should not be treated as a "central topic in cognitive psychology" (Kahneman, 1968) since there are several alternative and easier ways of reducing the signal-to-noise ratio in STVM.

The current data are equivocal with respect to the question of whether the form of buffer storage in the tachistoscopic recognition task is visual, or symbolic, or whether either store may be used. However, an experiment designed to distinguish between these alternatives has been carried out (Mitchell, in preparation) with results which support the Visual Buffer model.

References

AVERBACH, E. and CORIELL, A. C. (1961).   Short-term memory in vision.   *Bell System Technical Journal*, **40**, 309–28.

DOOST, R. and TURVEY, M. T. (1971).   Iconic memory and central processing capacity. *Perception and Psychophysics*, **9**, 269–74.

ERIKSEN, C. W. and COLLINS, J. F. (1967).   Some temporal characteristics of visual pattern perception.   *Journal of Experimental Psychology*, **74**, 476–84.

ERIKSEN, C. W. and COLLINS, J. F. (1968).   Sensory traces versus the psychological moment in the temporal organization of form.   *Journal of Experimental Psychology*, **77**, 376–82.

GIBSON, E. J. (1965).   Learning to read.   *Science*, **148**, 1066–72.

HABER, R. N. and Standing, L. G. (1969).   Direct measures of short-term visual storage. *Quarterly Journal of Experimental Psychology*, **21**, 43–54.

HABER, R. N. and STANDING, L. G. (1970).   Direct estimates of the apparent duration of a flash.   *Canadian Journal of Psychology*, **24**, 216–29.

KAHNEMAN, D. (1968).   Method, findings and theory in studies of backward masking. *Psychological Bulletin*, **70**, 404–25.

KROLL, N. E. A., PARKS, T., PARKINSON, S. R., BIEBER, S. L. and JOHNSON, A. (1970). Short-term memory while shadowing; recall of visually and of aurally presented letters. *Journal of Experimental Psychology*, **85**, 220–4.

LISS, P. (1968).   Does backward masking by visual noise stop stimulus processing?   *Perception and Psychophysics*, **4**, 328–30.

NEISSER, U. (1967).   *Cognitive Psychology*.   New York: Appleton-Century-Crofts.

PHILLIPS, W. A. and BADDELEY, A. D. (1971).   Reaction time and visual short-term memory.   *Psychonomic Science*, **22**, 73–4.

POSNER, M. I., BOIES, S. J., EICHELMAN, W. H. and TAYLOR, R. L. (1969).   Retention of visual and name codes of single letters.   *Journal of Experimental Psychology Monograph*, **79**, (1, part 2).

ROSS, B. M. (1969).   Sequential visual memory and the limited magic of the number seven. *Journal of Experimental Psychology*, **80**, 339–47.

SPERLING, G. (1960).   The information available in brief visual presentations.   *Psychological Monographs*, **74**, (11, Whole no. 498).

SPERLING, G. (1963).   A model for visual memory tasks.   *Human Factors*, **5**, 19–31.

SPERLING, G. (1967).   Successive approximations to a model for short-term memory. *Acta Psychologica*, **27**, 285–92.

TAYLOR, R. L. (1969).   Comparison of short-term memory and visual sensory analysis as sources of information.   *Journal of Experimental Psychology*, **81**, 515–22.

WARRINGTON, E. K. and SHALLICE, T. (1972).   Neuropsychological evidence of visual storage in short-term memory tasks.   *Quarterly Journal of Experimental Psychology*, **24**, 30–40.

REFRACTORINESS IN SAME–DIFFERENT JUDGMENTS

BARBARA J. RAEBURN† AND MICHAEL C. CORBALLIS

Department of Psychology, McGill University

Subjects judged as "same" or "different" pairs of visual stimuli, each either an X or an O, whose onsets were separated by intervals of 50, 100, 150 or 200 ms. When the interstimulus interval (ISI) was held constant within a block of 40 trials (Experiment I), decision time was independent of ISI. But when ISI was varied unpredictably from trial to trial (Experiments II and III), decision times increased the shorter the ISI. This effect was more marked for "same" than for "different" decisions. These results fail to support a single-channel interpretation, but suggest that variations in decision time depend on the subject's expectancy as to when the second stimulus will follow the first.

Introduction

If a subject must respond to two stimuli in succession, his reaction time to the second is typically delayed if the interval between their onsets is less than about 300 ms (Welford, 1952). The delay in reaction time, generally known as the "psychological refractory period" (after Telford, 1931), tends to increase approximately linearly as the interstimulus interval decreases. Smith (1967) has reviewed the evidence and contrasted various theoretical interpretations of it.

Davis (1964) and Corballis, Lieberman and Bindra (1968) have found similar delays in decision time when a subject must decide whether two rapidly presented visual stimuli are the same or different. This technique differs in one important respect from the double reaction time technique in that only a single response ("same" or "different") is required for each stimulus pair, so the delay cannot be attributed to response interference. Here, the "refractoriness" must be due either to stimulus interference or to some aspect of the central processing of information.

Davis (1964) favoured interpretation in terms of a limited-capacity, "single-channel" processing system (cf. Broadbent, 1958; Bertelson, 1966). Since the system can generally process only one stimulus at a time, he argued, a second stimulus closely following a first must be held in buffer storage until the processor is free. This waiting period causes the delay. However, if the stimuli are nearly simultaneous and if the subject knows this in advance, they may be grouped together and processed as a single unit, in which case there is little or no delay. Davis in fact found a *decrease* in decision latency for a 35-ms interstimulus interval compared with a 160-ms one under a condition in which the interval was held

†Present address: Department of Psychology, University of Auckland, Private Bag, Auckland, New Zealand.

constant within any block of trials. But latency was longest of all for the 35-ms interval when the intervals varied unpredictably from trial to trial.

Corballis *et al.* (1968) concluded rather differently. Their data did not yield any drop in decision latency at the shortest interstimulus interval, in this case 50 ms. The results were entirely consistent with the interpretation that delays in latency were due to stimulus masking. Unlike Davis, they presented their stimuli to the same spatial location, so that one would expect some masking to have occurred.

Another possible explanation for refractoriness, whether in double reaction time or in same–different experiments, appeals to the subject's expectancy (e.g. Adams, 1962; Elithorn and Lawrence, 1955; Smith, 1967). If the subject receives a range of different interstimulus intervals in random order, so the explanation goes, he will learn to "expect" the second stimulus somewhere around the mean interval. Hence latency should be fastest when the interval is in fact near the mean. It should be the slowest when the interval is shortest and the second stimulus is least expected. One might also expect slight delays at long intervals, although expectancy should not be so low as at correspondingly short intervals since the subject knows the stimulus must come sometime, given it has not already. Corballis *et al.* (1968) did find slight increases in latency with long interstimulus intervals. Expectancy differences associated with different interstimulus intervals should be minimal when the intervals are held constant within blocks of trials.

The present experiments, like Davis's (1964), investigate latencies to make same–different judgments both when interstimulus intervals are held constant over blocks of trials and when they are varied unpredictably from trial to trial. They are thus designed specifically to contrast "single-channel" and expectancy explanations of refractoriness in same–different decisions. The stimulus pairs were presented to different spatial locations, which should rule out masking effects.

Experiment I

In this experiment, subjects made same–different judgments about pairs of stimuli where interstimulus intervals were fixed within blocks of trials. Expectancy theory would predict that decision latency would be independent of interstimulus interval, whereas the single-channel theory would predict increasing latencies at short interstimulus intervals, perhaps with a drop at the very shortest intervals due to grouping.

Method

Subjects

Subjects were 24 paid students, ranging in age from 16 to 28 years. Mean age was 19·4 years.

Apparatus and procedure

The stimuli were presented on two multiple stimulus projectors (Grason-Stadler, Model E-4580-3) 3·8 cm wide by 5·1 cm high. The projectors were mounted side by side on a black screen, with their centres 4·1 cm apart. Each stimulus was either an X or an O.

The first stimulus always occurred on the left projector, the second stimulus on the right. The subject sat about 1·3 m from the screen, and viewed the projectors at eye level.

The response box consisted of two microswitch buttons horizontally placed 10 cm apart. Halfway between the two buttons was a "rest key". For half the subjects the button on the "preferred" (writing-hand) side of the response box was labelled "same"; for the other half it was labelled "different".

Order and timing of events in each trial were as follows: warning tone (300 ms); pause (500 ms); first stimulus (50 ms); interstimulus interval (variable); second stimulus (50 ms); intertrial interval (2000 ms).

Events in each trial were controlled by a multi-channel tape programmer. Reaction time was measured by a Hewlett Packard Electronic Counter, which was turned on immediately following the second stimulus, and turned off when the subject pressed a response button. If a response did not occur the counter was turned off by the onset of the next trial. A Hewlett Packard Digital Recorder printed out the time between onset and offset of the counter to the nearest ms and also recorded which button the subject pressed.

Interstimulus interval (ISI) was measured from the onset of the first stimulus to the onset of the second. Four different values were used: 50, 100, 150, and 200 ms. In this experiment, ISI was held constant during a block of trials, but was counterbalanced between blocks.

Each block consisted of eight practice trials, followed by 40 test trials. Order of presentation of each of the stimulus pairs (XX, OO, XO, OX) within the test trials was random, with the following exceptions: probability of occurrence of each stimulus pair was equal; each stimulus pair followed every other stimulus pair equally often; the largest possible run of repeated stimulus pairs (e.g. XX following XX) was three; the longest possible run of repeated response pairs (e.g. "same" following "same") was four.

Each subject was told to press the "same" button if the stimuli of the pair were the same (XX or OO), and the "different" button if the stimuli were different (XO or OX). He was told to use only the index finger of his preferred hand throughout the experiment, and to place his finger on the rest key between trials. The experimenter stressed the importance of responding quickly and accurately.

Results and Discussion

Errors and omissions made up less than 1% of the total response, as can be seen from Table I. They were excluded from further analysis.

TABLE I

Percentage of errors made at each interstimulus interval (ISI) in each Experiment

	ISI (ms)			
	50	100	150	200
Experiment I	0·7	0·8	0·6	1·2
Experiment II	0·6	1·5	0·6	2·5
Experiment III	0·2	1·0	0·6	1·5

Geometric means of the latency scores on correct trials were calculated for each subject for each judgment at each ISI. Geometric means were used to reduce the positive skew that is characteristic of reaction time data. A Latin Square analysis of variance was carried out on the means. Independent variables were ISI, judgment, order of ISI, blocks, and key position (i.e. "same" to preferred or non-preferred side).

"Same" latencies were significantly faster than "different" latencies ($F = 13·79$, $df = 1, 16$, $P < 0·01$). This is consistent with previous evidence using easily

discriminable visual stimuli (e.g. Corballis *et al.*, 1968; Egeth, 1966; Entus and Bindra, 1970; Nickerson, 1965). There were no other significant effects apart from an interaction between judgments and key positions ($F = 7.84$, $df = 1, 16$, $P < 0.05$).
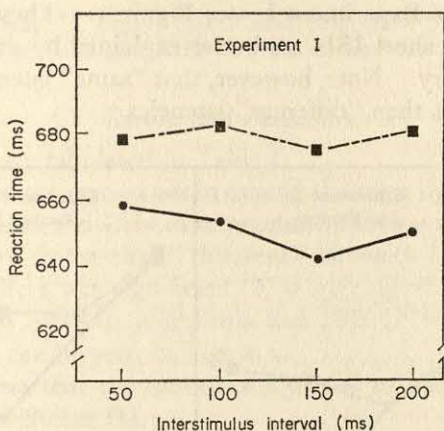
FIGURE 1. Mean reaction time as a function of ISI and judgment in Experiment I.
■ — — — ■ different; ●——● same.

Most important, however, was the fact that there was no evidence for any effect of ISI ($F < 1$, $df = 3, 48$). Figure 1 shows the means for each judgment at each ISI. This result confirms the conclusion of Corballis *et al.* (1968) that the ISI effect in *their* experiment was due to masking at short ISIs. It also throws doubt on single-channel theory (e.g. Davis, 1964).

Experiment II

This experiment differs from Experiment I only in that interstimulus intervals were varied from trial to trial, instead of remaining constant within blocks of trials.

Method

Subjects

The subjects were 12 paid students whose ages ranged from 16 to 24 years.

Apparatus and procedure

The apparatus was the same as in Experiment I.

The subjects were given a practice session of eight trials, followed by two test blocks of 80 trials each. ISIs (50, 100, 150 and 200 ms) were varied randomly within each block, except that each was presented an equal number of times. Other conditions were as in Experiment I.

Results and Discussion

Again, there were very few errors (see Table I).

Geometric means of latencies for correct responses were again computed for each judgment at each ISI for each subject, and analysis of variance carried out.

Again there was a significant difference between "same" and "different" latencies ($F = 17.24$, $df = 1$, 10, $P < 0.01$) with "same" the faster.

There was also a significant effect due to ISIs ($F = 17.34$, $df = 3$, 30, $P < 0.001$) and an interaction between ISIs and judgments ($F = 6.29$, $df = 3$, 30, $P < 0.01$). These effects are shown in Figure 2. These results contrast rather strikingly with those of Experiment I—see Figure 1. They suggest, indeed, that increases in latency at short ISIs are better explained by expectancy theory than by single-channel theory. Note, however, that "same" latencies show much more pronounced ISI effects than "different" latencies.
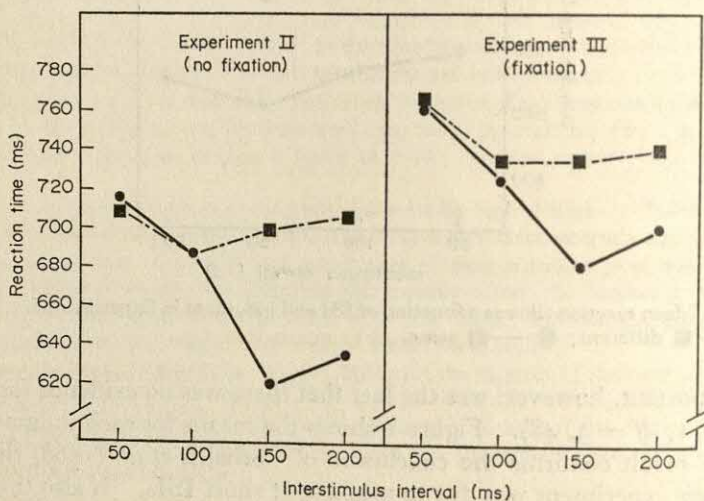
FIGURE 2.   Mean reaction time as a function of ISI and judgment in Experiments II and III.
■— — — — —■ different; ●——● same.

Another possible explanation for the results of this experiment concerns eye movements. The most efficient strategy at short ISIs might be to fixate between the two stimulus projectors, so that both can be perceived without eye movements. But at longer ISIs, a better strategy might be to look at left and right projectors in turn, since there would be time to shift fixation between stimulus presentations. In Experiment I the subjects could have adopted the most appropriate strategy for each block, since ISIs were constant within blocks. But in Experiment II the subject could never know which ISI would occur next. If he used a strategy of fixating each projector in turn, the time taken to shift fixation might conceivably delay perception of the second stimulus at short ISIs, and so increase decision latency.

Experiment III

This experiment was a replication of Experiment II except that the subjects were instructed to fixate a light placed exactly between the two projectors during each trial. While this might not completely eliminate eye movements, one would expect that any effects due to eye movements in Experiment II would be at least attenuated in this experiment.

Method

Subjects

The subjects were 12 student volunteers ranging in age from 19 to 25 years.

Apparatus and procedure

Exactly as in Experiment II, except that a small neon glow tube (0·04 W) was placed exactly between the two projectors, and the subjects were told to fixate it during each trial.

Results and Discussion

Errors, again rare, are tabulated in Table I.

Analysis of latencies was carried out in exactly the same way as in Experiment II. The effects of judgments and ISIs were practically the same as in Experiment II: "same" latencies were faster than "different" latencies ($F = 19·2$, $df = 1$, 10, $P < 0·01$), the ISI effect was significant ($F = 23·3$, $df = 3$, 30, $P < 0·001$) and so was the interaction between judgments and ISIs ($F = 5·8$, $df = 3$, 30, $P < 0·01$). These effects can be seen in Figure 2.

These results suggest that eye movements played no appreciable role in either Experiment II or Experiment III.

General Discussion

Our results do not really offer any support at all for single-channel theory, since Experiment I showed no effect of interstimulus interval. This is contrary to the results of experiments by Borger (1963), Creamer (1963) and Davis (1964), who did report evidence for refractoriness under fixed-interval conditions. However, the first two of these studies used double reaction-time paradigms, so the increases in the second reaction time at short intervals could have been due to response interference. This leaves Davis (1964), who did use a same–different paradigm. His study was actually similar to the present one; the main differences were that he used a wider range of intervals (35, 160, 600 and 810 ms), his stimuli were presented such that the second appeared below the first, and he used only two subjects in each of his fixed-interval and variable-interval conditions. It should also be noted that, in his fixed-interval condition, the only evidence that interval had any effect was a fairly small increase in latency for the 160-ms interval over the others.

Another feature of our results which is difficult to reconcile with single-channel theory is that the latency increases with short interstimulus intervals were much more pronounced for "same" than for "different" judgments in Experiments II and III. Single-channel theory attributes the increase to a delay in processing the second stimulus, so it should be independent of the nature of this stimulus. Our results show that it is not.

Posner and Boies (1971) have reached a similar conclusion. Their subjects, like ours, were required to decide as quickly as possible whether pairs of letters were the same or different, although they used a larger pool of letters than we did. By examining the influence on decision time of a competing auditory detection task, they concluded that the encoding of a letter does not require processing

capacity, but response selection does. Thus one would expect single-channel effects only when the subject must generate responses to *both* stimuli, as in the classic experiments on "psychological refractory period" (Welford, 1952).

Expectancy theory seems to provide the most adequate explanation of the results. It explains why interstimulus interval did not influence decision latency when the intervals were constant within blocks of trials (Experiment I). As we observed in the introduction, it also explains why, when the intervals varied unpredictably (Experiments II and III), there was a slight increase in latency at the longest interval as well as a sharp one at the two shortest intervals. However, Davis (1964) argued against an expectancy explanation of his results on the basis of a control experiment in which the first stimulus was simply a blank field, and the second either an X or an O. Subjects made a two-choice reaction to the second stimulus. There was some evidence that expectancy played a role under a variable-interval condition, but the effect was small compared with the effect of interstimulus interval under the variable-interval condition in the main experiment, where subjects made same-different judgments. Conceivably, however, subjects in the control experiment paid little attention to the blank stimulus, choosing perhaps to fixate the position in which the second stimulus was to appear. Thus expectancy might have played a more significant role in the same–different task, where subjects were forced to attend to the first stimulus.

We also require some explanation for why the effect of shortening the interval should have been so much more marked for "same" than for "different" judgments in Experiments II and III. A similar phenomenon has been reported by a number of other authors. Some (e.g. Bindra, Williams and Wise, 1965; Corballis *et al.*, 1968; Entus and Bindra, 1970) have suggested that "same" judgments are more influenced by variations in the discriminability of the stimuli than are "different" judgments. This interpretation can be reconciled with the expectancy-theory interpretation of our results if it is supposed that the second stimulus is imperfectly registered if it occurs at an unexpected moment, so that discriminability is worst when the stimulus is least expected.

On the other hand, it is not so easily reconciled with data reported by Posner and Boies (1971). They found an increasing divergence of "same" and "different" latencies in letter matching as the interstimulus interval increased from 0 to 500 ms, even though the intervals were constant within blocks of trials as in our Experiment I. The stimuli were presented to different spatial locations, so the effect cannot be attributed to masking, as in the experiments by Corballis *et al.* (1968). Posner and Boies suggest that "same" decisions are especially rapid when encoding of the first stimulus is optimal, which takes about 500 ms. However, this does not explain why we found no interstimulus interval effects at all in Experiment I, or why the "same" function appears to show an upward trend at only 200 ms in Experiments II and III. It is possible that encoding was more critical in Posner and Boies' experiments because their letters were drawn from an alphabet of 15 letters, whereas we used only the highly discriminable letters X and O.

We are left with the impression that many different factors can influence latencies when two stimuli must be processed in rapid succession. Single-channel effects

probably occur only when responses are required to both stimuli. When the stimuli must be matched in some way, other factors may be important. Corballis *et al.* (1968) have discussed the possible effects of masking, and Posner and Boies (1971) have documented the roles of preparation, encoding and central processing capacity. The present experiments suggest that expectancy may be yet another important influence.

References

ADAMS, J. A. (1962). Test of the hypothesis of psychological refractory period. *Journal of Experimental Psychology*, **64**, 280–7.

BERTELSON, P. (1966). Central intermittency twenty years later. *Quarterly Journal of Experimental Psychology*, **18**, 153–63.

BINDRA, D., WILLIAMS, J. A. and WISE, J. S. (1965). Judgments of sameness and difference: Experiments on decision time. *Science*, **150**, 1625–7.

BORGER, R. (1963). The refractory period and serial choice reactions. *Quarterly Journal of Experimental Psychology*, **15**, 1–12.

BROADBENT, D. E. (1958). *Perception and Communication.* London: Pergamon Press.

CORBALLIS, M. C., LIEBERMAN, W. and BINDRA, D. (1968). Discriminability and central intermittency in same–different judgments. *Quarterly Journal of Experimental Psychology*, **20**, 51–61.

CREAMER, L. R. (1963). Event uncertainty, psychological refractory period, and human data processing. *Journal of Experimental Psychology*, **66**, 187–94.

DAVIS, R. (1964). The combination of information from different sources. *Quarterly Journal of Experimental Psychology*, **16**, 332–9.

EGETH, H. E. (1966). Parallel vs. serial processes in multidimensional stimulus discrimination. *Perception and Psychophysics*, **1**, 245–52.

ELITHORN, A. and LAWRENCE, C. (1955). Central inhibition—some refractory observations. *Quarterly Journal of Experimental Psychology*, **7**, 46–127.

ENTUS, A. and BINDRA, D. (1970). Common features of "repetition" and "same–different" effects in reaction time experiments. *Perception and Psychophysics*, **7**, 143–8.

NICKERSON, R. S. (1965). Response times for "same"–"different" judgments. *Perceptual and Motor Skills*, **20**, 15–8.

POSNER, M. J. and BOIES, S. J. (1971). Components of attention. *Psychological Review*, **78**, 391–408.

SMITH, M. C. (1967). Theories of the psychological refractory period. *Psychological Bulletin*, **67**, 202–13.

TELFORD, C. W. (1931). Refractory phase of voluntary and associative responses. *Journal of Experimental Psychology*, **14**, 1–35.

WELFORD, A. T. (1952). The "psychological refractory period" and the timing of high speed performance: a review and a theory. *British Journal of Psychology*, **43**, 2–19.

SHORT-TERM MEMORY AND THE NATURE OF INTERFERENCE FROM CONCURRENT SHADOWING

NEAL E. A. KROLL

Department of Psychology, University of California, Davis, California

Trigrams were presented visually or auditorily and followed by a 12 s retention interval filled with shadowing numbers or letters. Auditory memory letters followed by letter shadowing were recalled less than auditory memory letters followed by number shadowing or visual memory letters followed by either type of shadowing. The latter three conditions did not differ among themselves. An analysis of the recall intrusions suggested that forgetting of auditory memory letters followed by letter shadowing was caused mainly by a confusion between covert rehearsals and shadowing activity, while forgetting in the other three conditions was caused primarily by proactive interference from earlier memory trials.

Introduction

Apparently adult humans typically use verbal memory codes even if the memory stimulus is presented visually, at least if a verbal code is readily available (e.g. Conrad, 1964; Sperling, 1967). This verbal code is maintained at the expense of the potential visual code which is often used at maximal efficiency for only a few seconds after presentation (Posner, Boies, Eichelman and Taylor, 1969).

Shadowing letters of the alphabet during the retention interval reduces the recall of auditorily presented memory letters more than the recall of visually presented memory letters (Kroll, Parks, Parkinson, Bieber and Johnson, 1970; Parkinson, 1972). The shadowing task appears to have its deleterious effect on the memory of auditory stimuli at least partially through its interference with effective subvocal rehearsal. Thus the visual stimuli are not, under these conditions, recoded into verbal memory codes (Parkinson, Parks and Kroll, 1971; Salzberg, Parks, Kroll and Parkinson, 1971) but are, rather, retained via visual memory codes (Parks, Kroll, Salzberg and Parkinson, 1972). The question remains as to how the shadowing task interferes with subvocal rehearsal. At least two, not necessarily mutually exclusive, possibilities exist. One is that the shadowing task reduces the number of subvocal rehearsals possible while not affecting the amount of visual imagery occurring during the retention interval. Another possibility is that the subject's subvocal rehearsals are similar to his vocalizations required by the shadowing task and become confused with them; while visual imagery, being qualitatively different, remains relatively free of contamination from the shadowed letters. The latter seems to be the more important source of interference since subjects asked

to press a button every time they "think of the memory letters in any way" report nearly as many rehearsals following auditory as following visual memory letter presentation, but the effectiveness of each rehearsal seems to be much less following auditory presentation (Kroll and Kellicutt, 1972).

The present experiment attempts to evaluate these two possibilities by means of varying the type of materials shadowed. Specifically, while the memory stimuli are always three alphabetic letters, subjects shadow letters on half of the trials and shadow numbers on the other half. Assuming that shadowing numbers is roughly equal in difficulty to shadowing letters, the two tasks should allow approximately the same number of covert rehearsals, but it should be easier to differentiate these rehearsals from the shadowed numbers than from the shadowed letters. Thus, if the first hypothesis is correct, the two shadowing tasks should generate equal interference, resulting in the recall of the auditory memory letters being equally bad (as compared with visual memory letters). Alternatively, if the second hypothesis is correct, shadowing numbers should cause much less interference, resulting in the recall of auditory memory letters being much better with number shadowing. In either case, if the visual memory letters are, in fact, rehearsed via visual imagery, recall of the visual memory letters should not be much affected by the type of material shadowed.

Method

Subjects

Twenty students (11 females and 9 males) from introductory psychology courses at the University of California, Davis, served as subjects and received course credit for their participation. One additional student was not used because of an inability to learn how to shadow proficiently within the allotted time.

Materials and apparatus

Shadow lists were constructed by selecting 35 letters (from the set A, C, D, I, M, O, R, S, U, and Y) or 35 numbers (1–10) at random, with the use of a random number table and with the restriction that no letter or number immediately follow itself. A female voice was then recorded reading these lists at a rate of two letters or numbers per s. Memory stimuli were trigrams selected from the letters B, F, G, H, J, K, L, N, Q, T, X, and Z such that no letter was repeated in the trigram and that each trigram had an association value equal to or less than 29% (Witmer, 1935). Assignment of memory trigrams to trials was such that the same letter would not appear in the trigrams of two adjacent trials.

On trials where the memory trigram was presented auditorily, three of the female-voiced letters or numbers were removed and replaced by a male voice reading the memory trigram. The letters of the memory trigram replaced the ninth, tenth, and eleventh items of the list: thereby resulting in a 4-s lead-in to the memory trigram and a 12-s retention interval. The subject was instructed to shadow both the female- and male-voice letters, but to also remember the male-voice letters.

On trials where the memory trigram was presented visually, the three letters were flashed individually at a 2/s rate, each with an on-time of approximately 200 ms, contiguous with the female voice reading the ninth, tenth, and eleventh items in the shadow list. The visual letters (8·5 cm in height) appeared via a single one-plane readout unit (Industrial Electronic Engineers, Series 80000) located approximately 1·8 m from the subject and approximately at eye level. The subjects were instructed to continue shadowing the female voice while they watched and remembered the visually presented letters.

Lists to be shadowed and auditory memory letters were presented binaurally via earphones (Superex, ST-Pro B) from one track of a stereo tape recorder. Signals on the other track

controlled visual presentations by means of a voice-sensitive relay. Sequencing of the visual letters was accomplished by means of a card punch (IBM, model 026) and associated relay and timing circuitry.

Procedure

Each subject reported individually and was given shadowing practice similar to that given in earlier experiments (e.g. Kroll *et al.*, 1970), except that in the present experiment half of the practice lists were composed of numbers. After the subject had learned to shadow with at least 80% accuracy and not to leave long pauses in his shadowing, he was given memory instructions followed by the memory trials. Each subject received two blocks of 22 trials each, with a rest break of approximately 5 min between blocks. In each block, 11 number shadowing and 11 letter shadowing trials were interspersed. Half of the subjects received auditory memory letters in their first block and visual in their second, while the other half of the subjects received visual memory letters in their first block. While each subject received a particular memory trigram only once, across all subjects each trigram appeared equally often under each of the four experimental conditions. There was a 14-s inter-trial interval, during which the subject wrote his recall response and covered it so that he could never see his earlier responses.

Results and Discussion

Since interpretation of the recall data depends largely on the validity of the assumption that numbers and letters are, roughly, equally difficult to shadow, shadow performance will be discussed first. While most subjects reported that the numbers seemed somewhat easier to shadow, the actual shadow performance of a randomly selected 10 subjects showed little actual difference in performance (18·1% vs. 15·4% errors during the retention interval for shadowing letters vs. numbers respectively). Since the difference is slight and since the recall of visual memory letters, known to be sensitive to shadow list difficulty (Parkinson *et al.*, 1971), was not much affected (see below), the assumption of a rough equivalence seems justified.

Table I presents the mean percentage of correct recall responses for visual and auditory memory letters after shadowing letters or numbers. The first two trials of each block (one each with letter and number shadowing) were not included in these summaries and analyses. In Free Recall scoring, each letter of the trigram was scored individually, regardless of order. In Perfect Recall scoring, the entire trigram had to be recalled, in proper order, to be scored correct.

TABLE I

Mean percentage of correct recall as a function of memory letter presentation modality, shadow task, and scoring procedure

Scoring procedure: Presentation mode:	Free recall		Perfect recall	
	Visual	Auditory	Visual	Auditory
Shadowing letters:	66·7	50·6	27·5	17·0
Shadowing numbers:	71·9	72·6	28·5	29·5

According to analyses of variance (a one-way, four-level, within-subject analysis for each scoring procedure), there was significant variation among the conditions for both Free Recall and Perfect Recall scoring procedures ($F = 20\cdot14$, $df = 3, 57$, $P < 0\cdot001$; and $F = 4\cdot929$, $df = 3, 57$, $P < 0\cdot005$ respectively). Tukey's test of Honestly Significant Differences (HSD) suggests that recall of the auditory memory letters with letter shadowing was significantly poorer than the other three conditions which did not differ significantly among themselves. This is true both for Free Recall scoring, where the HSD corresponding to $P < 0\cdot01$ two-tailed is $10\cdot58$, and for Perfect Recall scoring, wherethe HSD corresponding to $P < 0\cdot05$ two-tailed is $9\cdot81$. Thus the per cent correct response data support the hypothesis that shadow interference is largely a result of the subject confusing the shadow letters with his subvocal rehearsals of the auditory memory letters.

There are two precautions that should be observed when attempting to compare the absolute levels of recall given in Table I. The first is that since auditory memory is less disturbed by number shadowing than by letter shadowing, subjects may recode visual memory letters into verbal memory codes on trials with number shadowing. The slight superiority of visual presentation may be the result of the subjects occasionally both retaining the visual image and, in addition, forming a verbal memory code. The second precaution to be considered is that visual presentation involves concurrent shadowing of extraneous material while auditory presentation does not. This may have limited the attention given visually presented memory trigrams, resulting in a low estimate of visual memory. However, the main point is not that the memory of auditory letters with number shadowing is as good as the memory of visual letters, but rather that the memory of auditory memory letters is affected by the type of material being shadowed much more than is the memory of visual letters.

The present experiment allows a further test of the hypothesis that shadow interference is largely a result of a confusion between subvocal rehearsals and shadowing activity. Since only 10 letters were ever used as shadow letters, a comparison of the percentage of intrusions from the shadow letters can be compared across conditions. [This type of intrusion analysis was attempted by Salzberg (1971). While highly suggestive of the same type of conclusions reached in the present study, the overlap of letter sets which were potential contributors to proactive and retroactive interference made interpretation difficult and consequently it was omitted from the published version (Salzberg *et al.*, 1971).] If one of the letters from the 10 used as shadow letters is given as a memory letter on a trial with number shadowing, it is most likely a random guess. However, if subjects, when shadowing letters, confuse the shadow letters with their covert rehearsals, there should be many more intrusions from these 10 letters on trials with auditory memory letters and letter shadowing. Table II presents the proportion of intrusions coming (a) from memory letters other than those three presented on that trial (a set of 9 letters), (b) from shadow letters (a set of 10 letters), and (c) from other letters (potentially a set of 4 letters, although no subject ever gave "W" in a recall response, plus those letters from the memory letter set that the subject had not, by the time of the intrusion, experienced as a memory letter) as a function of the type of trial on which the intrusion occurred. Omissions were not included

and if an intrusion letter was given more than once on a particular trial (e.g. "HHF" or "HFH") it was counted only once. As predicted by the hypothesis, the trials with auditory memory letters and letter shadowing had, by far, the highest proportion of shadow letter intrusions.

TABLE II

Proportion of intrusions from shadow letters, memory letters, and other letters as a function of memory letter presentation modality and shadow task

| Shadow task: | Letters | | Numbers | |
Presentation mode:	Visual	Auditory	Visual	Auditory
Shadow letters:	0·392	0·771	0·105	0·314
Memory letters:	0·580	0·195	0·850	0·509
Other letters:	0·028	0·034	0·045	0·176

The relatively high proportion of intrusions from earlier experienced memory letters in the other three conditions suggests that a long-term memory process is playing a large role in this type of task and that this long-term memory process suffers proactive interference from earlier trials. A rough estimate of the effect of the build-up of proactive interference can be found by looking at the mean percentage of correct recall responses on the first trial of each block (i.e. the subject's first visual and first auditory memory trigram). The mean percentage of correct recall (Free Recall scoring) on the first visual trial was 83·4% with number shadowing and 96·7% with letter shadowing; while it was 76·7% on the first auditory trial with number shadowing and 66·7% with letter shadowing. It should be emphasized that these percentages are based on very few data (10 subjects, one trial each, for each percentage), and one should probably not try to judge the relative amount of proactive interference in each of the tasks from these figures. However, they do suggest a superiority on the first trial, even though the memory trigrams on the first trial were of a lower association value than those used on later trials (an average of 6% vs. an average of 16·96%).

Thus recall of memory letters after a retention interval filled with shadowing activity seems to involve a long-term memory process which is subject to proactive interference from earlier memory trials. However, subjects are able to partially counter this proactive interference by rehearsing the current memory trigram during the shadow-filled retention interval (Kroll and Kellicutt, 1972). The rehearsal of the memory stimulus during shadowing appears to take the form of a visual image following visual presentation (Parks *et al.*, 1972) and the form of an auditory image and/or a subvocal response following auditory presentation (Parkinson *et al.*, 1971; Salzberg *et al.*, 1971). The present experiment suggests that letter shadowing reduces the effectiveness of the rehearsal following auditory memory letters through a confusion between memory letter rehearsal and shadow letters.

References

CONRAD, R. (1964). Acoustic confusions in immediate memory. *British Journal of Psychology*, **55**, 75–83.

KROLL, N. E. A. and KELLICUTT, M. H. (1972). Correct recall in short-term memory as a function of covert rehearsal and of intervening task. *Journal of Verbal Learning and Verbal Behavior*, **11**, 196–204.

KROLL, N. E. A., PARKS, T., PARKINSON, S. R., BIEBER, S. L. and JOHNSON, A. L. (1970). Short-term memory while shadowing: recall of visually and of aurally presented letters. *Journal of Experimental Psychology*, **85**, 220–4.

PARKINSON, S. R. (1972). Short-term memory while shadowing: multiple item recall of visually and of aurally presented letters. *Journal of Experimental Psychology*, **92**, 256–65.

PARKINSON, S. R., PARKS, T. E. and KROLL, N. E. A. (1971). Visual and auditory short-term memory: effects of phonemically similar auditory shadow material during the retention interval. *Journal of Experimental Psychology*, **87**, 274–80.

PARKS, T. E., KROLL, N. E. A., SALZBERG, P. M. and PARKINSON, S. R. (1972). Persistence of visual memory as indicated by decision time in a matching task. *Journal of Experimental Psychology*, **92**, 437–8.

POSNER, M. I., BOIES, S. J., EICHELMAN, W. H. and TAYLOR, R. L. (1969). Retention of visual and name codes of single letters. *Journal of Experimental Psychology Monograph*, **79** (No. 1, Pt. 2).

SALZBERG, P. M. (1971). Visual and auditory short-term memory: the differential effects of acoustically similar shadow material. Master's thesis, Sacramento State College, Sacramento, California.

SALZBERG, P. M., PARKS, T. E., KROLL, N. E. A. and PARKINSON, S. R. (1971). Retroactive effects of phonemic similarity on short-term recall of visual and auditory stimuli. *Journal of Experimental Psychology*, **91**, 43–6.

SPERLING, G. (1967). Successive approximations to a model for short-term memory. In SANDERS, A. F. (Ed.), *Attention and Performance*. Amsterdam: North-Holland Publishing Company. (A special edition of *Acta Psychologica*, **27**, 285–92.)

WITMER, L. R. (1935). The association value of three-place consonant syllables. *Journal of Genetic Psychology*, **47**, 337–59.

CODING PROCESSES IN THE FREE RECALL OF ASSOCIATED WORD LISTS

W. A. MATTHEWS AND A. J. G. WARING

Department of Psychology, University of Southampton

An experiment is reported on the free recall of lists of 18 words. The lists consisted of words that were strongly or weakly associated to the eliciting stimulus, were members of either a single category or more than one category and were arranged into sets of three associated words or six associated words. Each list was presented once only and free recall required after a filled interval of 18 s. One half of the subjects were provided with cues at recall: the remainder were given no assistance. Recall efficiency increased with association level and with restricted category membership of the words in a set. Clustering at recall (when adjusted for opportunity) was greater in the 6 × 3 lists than in the 3 × 6 lists. The provision of cues only led to a significant facilitation in recall with lists consisting of 6 sets of 3 words. A detailed analysis of the recall data suggested the importance of the category membership variable in the production of the later recall.

Introduction

Among a number of variables shown to affect the efficiency of free recall is that of associative frequency. Lists composed of sets of words strongly associated to eliciting stimuli are more efficiently recalled than lists of weak associates to the stimuli or control lists of non-associated words (e.g. Jenkins and Russell, 1952; Matthews, 1966). However, the associative frequency measure (i.e. how often the response is given to a presented stimulus) ignores the categorical relationships between words within associated lists and these may be manipulated as well as the associative measures. Thus an associated set can be conceptually unified (or homogeneous) in the way that the associates to "body"—arm, face, leg, nose—are clearly parts of the original stimulus and can be conceptualized as "body (parts of)", or the associative group can be conceptually divergent (or heterogeneous) where associates to bread—crumb, flour, knife, white—though clearly associated to bread do not in themselves form a conceptually unified category. An experiment by Marshall, manipulating word pair interrelationships, cited in Cofer (1965), suggests that these two factors can be separately manipulated.

This experiment attempted this type of manipulation but the problem which is raised by such operations is that although the associative strengths of the associates to the eliciting stimuli may be equalized in the homogeneous and heterogeneous sets, the mean strength of association of the homogeneous lists will be

greater when the total number of associative connections in a set is taken into consideration. The associative interconnections of the items in the homogeneous lists, being categorically related, will be greater than the strength of the interconnections in the heterogeneous lists. Thus, on an inter-item associative index measure such as the one used by Deese (1959), the two types of lists will differ. It might therefore be expected that high-association homogeneous lists would show superior recall to high-association heterogeneous lists and that a similar (and possibly greater) effect would occur at low association levels. An economical argument based on a simple associationistic model of storage and recall would produce similar predictions about the relative efficiency of recall to those produced by a model based on a "cognitive" coding by category membership. In the latter it might be argued that associative homogeneity would make a categorization process at input simple and effective, thus facilitating storage and retrieval. In the former, the overall stronger association level should similarly facilitate recall. Cofer (1966) suggests that both processes may operate but in different situations. However, the detail of the recall protocols might enable a distinction to be made between these alternative views so that the particular processes operating in this situation can be defined. If homogeneity facilitates category coding, recall might show more recalled clusters preceded by the eliciting stimulus than in the heterogeneous condition, whereas stronger inter-item associative connections might be expected to lead to some flexibility in the order of report within recalled word sets. The types of error might also differ.

A number of experiments in the area of free recall have been interpreted as indicating the operation of the cognitive type of category coding process in which the subjects impose category labels on the presented items and store these labels at one level and the constituent items at another. Recall of the category, or coding, label is thought to be a prerequisite for the recall of the items (cf. Cohen, 1966; Segal, 1969), although, with the imprecise definition given to "cognitive" coding, recall facilitation when cues are available or are presented still allows the use of associationistic models to explain the effects. Generally, a fair amount of evidence exists which indicates that cues have to be present both at presentation and at recall; that there has to be no ambiguity about category membership; and that there has to be a reasonable number of categories in the list, so that some category forgetting is possible, if the provision of cues is to show facilitation in recall (Tulving and Pearlstone, 1966; Tulving and Osler, 1968; Wood, 1969).

The aims of this experiment were:

(a) to investigate whether cueing facilitates recall and whether its effects, if any, vary as a function of list structure and list type;

(b) to investigate whether conceptually homogeneous lists (i.e., lists with a tighter intra-set structure and a higher intra-set associative strength) show facilitation in recall over that found in heterogeneous lists of similar levels of direct associative strength, to ascertain whether this varies as a function of list structure and list type, and to investigate the details of such recall, e.g. order effects, error data.

TABLE I
Sample of lists used in experiment

	3 × 6				6 × 3		
HA	HA	LA	LA	HA	HA	LA	LA
Homogeneous	Heterogeneous	Homogeneous	Heterogeneous	Homogeneous	Heterogeneous	Homogeneous	Heterogeneous
Laugh	Action	Govern	Plough	Big	Life	Rain	Wonder
Clown	Battle	Authority	Apples	Enormous	Happiness	Leaves	Art
Funny	Dance	Direct	Crows	Tall	Death	Trees	Guess
Comedy	Sport	Impose	Narrow	Cold	Cabin	Eat	River
Amusing	Movement	Minister	Path	Ice	Wood	Pleasure	Clear
Joke	Speed	Subject	Team	Snow	Ship	Enjoy	Path
Night	Bread	Dramatist	Disease	Dress	Jungle	Food	Poor
Tired	Crumb	Emotion	Effect	Skirt	Hot	Health	Box
Sleep	Food	Feeling	Cuts	Suit	Forest	Sick	Thin
Relax	Knife	Nerves	Mind	House	Boy	Beast	Family
Slumber	Flour	Agony	Temperature	Roof	Cap	Field	Age
Rest	White	Expression	Waste	Wall	Friend	Grass	Distant
Body	Mountain	Ink	Forget	Fright	Horse	Cloth	Use
Arm	Snow	Drip	Cry	Terror	Mare	Piece	Action
Face	Green	Blob	Date	Fear	Jockey	Rolls	Wear
Leg	Climb	Spot	Keys	Language	Joke	Front	Article
Nose	Forest	Leak	Realize	Speech	Practical	Row	Lost
Shoulder	Rock	Mark	Never	Talk	Smile	Rank	Soiled

Materials and Method

List construction

Sixteen 18-word lists were constructed from continued word association norms (Matthews, 1969) to give two examples of each of the following four types of list: high-association conceptually homogeneous lists (HA Hom), high-association conceptually heterogeneous lists (HA Het), low-association conceptually homogeneous lists (LA Hom), and low-association conceptually heterogeneous lists (LA Het). Two structures were used with each of these types of list: three sets of six words and six sets of three words. Each set consisted of a core word, which was used in the generation of the original association norms, followed by either 2 or 5 of the associations given to it. HA sets contained associations to a stimulus given by at least 20 subjects from the norm-producing group of 50 subjects, and LA sets contained associations given by at least one, but fewer than three, subjects. Homogeneous sets were constructed in such a way that the associations were synonymous or fell within a single, easily defined category so far as the available norms allowed; the heterogeneous sets contained words which were associations of the appropriate strength to the core word but which were not synonyms nor had clear associations to a single superordinate category. Thus, using an associative measure based on the relationships of the words within a set to an eliciting stimulus, the homogeneous and heterogeneous lists were equated. On a measure which takes into account the complete range of interword relationships within an associated set (e.g. Deese's IIAI), the homogeneous and heterogeneous conditions will differ, with the homogeneous sets (and lists) showing stronger inter-item and intra-set associative relationships. An empirical confirmation of these differences was obtained using both a rating task and a sorting task. (See Table I for a sample set.) Due to the limitations of the material, two words were duplicated in the lists.

List validation

In a rating task, eight undergraduate students were asked to indicate which of the words in each list were "related" i.e., associated. The words were printed in a random order on a sheet of paper in a half 18×18 matrix such that each word could be compared with each other. Subjects indicated which words were related by entering a tick in the appropriate cell. The total number of marked relationships was entered into the appropriate cell of a master matrix.

In the sorting task, after an interval of one week, the same eight subjects were asked to card-sort the same lists. Each word in the list was printed on a 10.2×7.6 cm card and each list was given to the subject as a pack which he had to sort into "sets of words which you feel are related to each other in some way". No specification was given as to the number of sets which would be needed or the number of words in each set. The words sorted by a subject into a set were deemed to be related to each other and for each such interrelation a unit score was entered into a matrix equivalent to that used for the rating task.

Both total matrices, for each list, were then scored as follows:

(a) within category associations—the average was taken of all the values of the cells appropriate to the experimenter-designated sets. Thus in a six-word set there are 15 interrelations, i.e., 15 cells; in the case of three-word sets there are three interrelations. The value for the list was obtained from the mean of the set values;

(b) between category associations—the totals from all the residual cells, not used in (a), were averaged.

Both methods give similar results and the mean scores for measures (a) and (b) derived from the rating task only are given in Table III, Rows H and I, respectively. On the basis of the scoring system used, these show that for the homogeneous and heterogeneous pairs of lists at each list structure and associative level, the homogeneous lists were seen by these subjects as containing more tightly conceptually unified sets than the corresponding heterogeneous lists.

Experimental design

Associative frequency (high and low), list structure (three groups of six words and six groups of three words), the homogeneity/heterogeneity of the items within an associative set and the presence or absence of cues at recall were the manipulated variables. The first three factors and the combinations of them were within-subject variables; the last was a between-subject variable. A single presentation of each list type was given but to facilitate the identification of the relevant categories this was "blocked" in that the members of a set were presented sequentially rather than being randomized among the other members of the list (cf. Bower, Clark, Lesgold and Winzenz, 1969). To reduce recency effects in recall and possibly ensure sampling from secondary memory, a counting task (unrelated to the learning process) was interpolated between the end of presentation and the beginning of free recall.

Procedure

The lists were recorded on magnetic tape and presented auditorily at a rate of 1 word/s. Each list was presented only once. An interpolated task was inserted between the end of each list and the beginning of its recall, which consisted of the subject reading aloud, in unison with the rest of the test group, a list of numbers which had been provided. To synchronize reading, a series of pulses from an oscillator was recorded on the tape, beginning 0·75 s after the last word of the list and continuing at the rate of 1 per 0·75 s for a total of 18 s. A different sequence of digits was used after each list so that the subjects had to attend to the reading task. Written free recall, for which 2 min were allowed, began immediately after the last number had been spoken. It was stressed that the words could be written down in any order. So that the subjects were familiar with the requirements of the task, two practice lists, composed of pronounceable nonsense syllables, preceded the word lists. To enable the experimental manipulations to be achieved, each subject was provided with a response pad with response sheets interleaved with the digit sheets. The subjects in the cued recall condition were each provided with separate cue sheets for each list presented. These were interleaved with the digit sheets. In the case of the 6 × 3 lists, the cue sheets had the six core words printed on them; for the 3 × 6 lists, the given words were the three core words plus one association from each of the groups. In place of the cue sheets, subjects in the non-cued recall condition had blank pieces of paper interleaved with the digit sheets. Thus the recalled words, plus the cue words where appropriate, were available throughout the recall interval. At the end of recall the response sheets were turned over and not referred to further.

Three different randomizations of the word sets within the lists and of the words within the sets were used. In each of these orders the core word of each set preceded the rest of the members of that set. The three randomizations corresponded to three separate groups of subjects.

Subjects

The subjects were 80 men and women undergraduates who attended the experiment in three separate groups. Within each group, half of the subjects were randomly allocated to the cued recall condition. The remainder were given non-cued recall. One subject from the cued recall group was discarded due to a failure to score more than zero in seven of the lists. One subject (chosen at random) was discarded from the non-cued condition to equate subject numbers in analysis 1. This subject was replaced in the data reported in analysis 2.

Results

Analysis 1

Interest in the first analysis was centred on the differences between cued and non-cued recall. The recall scores which were used excluded the cue words which had been recalled by the subjects in each condition. These scores are shown in

Table II. This indicated that recall of HA lists was better than recall of LA lists ($F = 257 \cdot 37$, $df = 1$, 76, $P < 0 \cdot 01$); recall of homogeneous lists was better than recall of heterogeneous lists ($F = 163 \cdot 63$, $df = 1$, 76, $P < 0 \cdot 01$). Differences

TABLE II

(a) *Total correct recall less cue words recalled*

	HA				LA			
	3 × 6		6 × 3		3 × 6		6 × 3	
	HOM	HET	HOM	HET	HOM	HET	HOM	HET
Cued	500	359	467	421	363	272	379	265
Non-cued	478	364	375	378	344	272	345	240
Total	978	723	842	799	707	544	724	505

39 subjects contributed to each total in the cued recall condition and 39 different subjects contributed to the scores in non-cued condition.

(b) *Total scores for conditions less cue words recalled*

	HA	LA	HOM	HET	6 × 3	3 × 6	
Cued	1747	1279	1709	1317	1532	1494	3026
Non-cued	1595	1200	1542	1254	1338	1458	2796
Total	3342	2479	3251	2571	2870	2952	

due to list structure were not significant but there was a significant interaction with homogeneity ($F = 10 \cdot 13$, $df = 1$, 76, $P < 0 \cdot 01$). A test of simple main effects (Kirk, 1968) and inspection of Table II(a) indicates that in the homogeneous condition recall is better in 3 × 6 lists than in 6 × 3 lists ($F = 34 \cdot 6$, $df = 1,152$, $P < 0 \cdot 01$) while in the heterogeneous condition recall is no better in 6 × 3 lists than in 3 × 6 lists. There is also a significant three-way interaction between list structure, association level and homogeneity ($F = 27 \cdot 4$, $df = 1$, 76, $P < 0 \cdot 01$). Homogeneous scores are higher in 3 × 6 lists than in 6 × 3 lists ($F = 22 \cdot 93$, $df = 1$, 152, $P < 0 \cdot 01$) in the HA condition while there is no difference in the LA condition; heterogeneous scores reverse this trend with scores being higher in 6 × 3 lists than in 3 × 6 in HA conditions ($F = 7 \cdot 16$, $df = 1$, 152, $P < 0 \cdot 05$) and there is again no difference in the LA conditions. Cued recall was not significantly different from non-cued recall but the interaction between cued/non-cued recall and list structure was significant ($F = 6 \cdot 03$, $df = 1$, 76, $P < 0 \cdot 05$). Table II(b) and a test of simple main effects shows that the provision of cues had no facilitating effect on 3 × 6 lists but clearly facilitated the recall of 6 × 3 lists ($F = 21 \cdot 75$, $df = 1$, 152, $P < 0 \cdot 01$). Further inspection of the totals suggests that there is some facilitation from cues in the HA and homogeneous conditions, which is where the effects of providing a cue should be most useful, but neither of the appropriate interactions (cue × association level; cue × homogeneity) reaches the 0·05 level of significance.

Analysis 2

This analysis was carried out on the total recall scores of 40 subjects in the non-cued recall condition (with recalled cue words included) and was concerned with the general characteristics of the recall data. The scores are shown in Table III, Row A.

TABLE III

Recall data from 40 subjects in the non-cued recall condition

| | HA | | | | LA | | | |
| | 3 × 6 | | 6 × 3 | | 3 × 6 | | 6 × 3 | |
	HOM	HET	HOM	HET	HOM	HET	HOM	HET
A	797	598	590	611	586	470	541	389
B	217	205	270	318	191	193	282	238
C	3·67	2·92	2·18	1·92	3·07	2·44	1·91	1·63
D	153	132	204	240	134	125	187	143
E	19·2	21·8	34·5	39·4	28·7	27·0	34·4	36·8
F	46·3	34·2	56·4	50·6	66·7	44·1	50·2	40·1
G	0·45	0·28	0·52	0·50	0·35	0·21	0·45	0·26
H	6·6	4·1	7·1	4·8	4·2	1·6	4·7	2·2
I	0·31	0·78	0·40	0·60	0·23	0·47	0·64	0·84

A=total number of items recalled, B=number of sets recalled, C=average number of items recalled per set, D=number of set labels recalled, E=set label word recall stated as a percentage of total recall for the condition, F=the percentage of recalled clusters which were preceded by the appropriate core word, G=mean clustering scores calculated from Dalrymple-Alford formula (1970), H=mean within set association values from rating task, I=mean between set association values from rating task.

In addition, the recall of the words presented at each list position was calculated for each list type although the detailed scores are not given. These data show a general primacy effect and, despite the interpolated task, a slight recency effect in the total recall of the list. The recall of the words within the associated sets showed a similar pattern. Homogeneity seems to flatten these curves, so reducing the more usual serial position effects within sets. Recall of the core word for each associative set tends to be higher than the recall of the rest of the items in the group (Table III, Row E) (cf. Segal, 1969, who reported a similar effect).

Clustering scores (C) are shown in Table III, Row G. These were computed according to the formula developed by Dalrymple-Alford (1970),

$$C = \frac{R - CC}{\text{Max} - \text{Min}},$$

where R is the number of repetitions in the protocol, CC is the estimated chance level of repetitions for that protocol, as formulated by Bousfield and Bousfield (1966) ($CC = (\Sigma k^2/n) - 1$), where k is the number of items recalled in a set. Max is the maximum number of repetitions which could be produced from the words in the protocol ($n - K$, where n is the number of words and K is the number of sets recalled), and Min is the minimum number of repetitions which are necessarily produced from the recalled words (Min = $2NL - n - 1$; where NL is the number of words recalled in the best recalled set).

A four-way analysis of variance was performed on these scores. This indicated that HA lists showed more clustering than LA lists ($F = 49.94$, $df = 1$, 39, $P < 0.001$); homogeneous lists showed more clustering than heterogeneous lists ($F = 64.27$, $df = 1$, 39, $P < 0.001$) and 6×3 lists showed more clustering than 3×6 lists ($F = 48.06$, $df = 1$, 39, $P < 0.001$). Subject variance was also significant ($F = 2.71$, $df = 39$, 39, $P < 0.005$). There was a significant interaction between list structure and association level ($F = 6.4$, $df = 1$, 39, $P < 0.05$). Inspection of Table III, Row G, and a test of simple main effects indicates that in high-association lists clustering is higher in 6×3 lists than in 3×6 lists ($F = 19.38$, $df = 1$, 39, $P < 0.01$) but the differences in the low-association condition are not significant. This is related to the three-way interaction between list structure, association level and homogeneity ($F = 7.87$, $df = 1$, 39, $P < 0.01$). Table III, Row G, and a test of simple main effects shows that in the high-association conditions clustering is higher in 3×6 homogeneous lists than in 3×6 heterogeneous lists ($F = 5.25$, $df = 1$, 78, $P < 0.01$), while there is no significant difference between clustering in 6×3 homogeneous and heterogeneous lists. The reverse of this occurs in the low-association conditions with higher clustering in 6×3 homogeneous lists than in 6×3 heterogeneous lists ($F = 6.56$, $df = 1$, 78, $P < 0.05$), there being no significant difference in clustering between 3×6 homogeneous and 3×6 heterogeneous lists.

The mean rate of intrusions is 4.0 items per subjects for intra-experimental intrusions (I) and 10.2 items for extra-experimental intrusions (E). These values are shown in more detail in Table IV. I intrusions are words which had been presented in previous lists; E intrusions had not been presented to the subject. Heterogeneous lists contain more I but fewer E intrusions than homogeneous lists. Repetition errors were not separately listed as only four occurred in the experiment. Most of the intrusions occur in the last half of the subjects' recall protocol for each list (134 first half, 414 second half). This could imply that subjects guessed to extend their recall to a "reasonable" length for each list in the absence of further known correct responses.

TABLE IV

Intrusion errors by condition shown by 40 subjects in the non-cued recall condition

| | HA | | | | LA | | | |
| | 3 × 6 | | 6 × 3 | | 3 × 6 | | 6 × 3 | |
	HOM	HET	HOM	HET	HOM	HET	HOM	HET
I	7 (0.12)	14 (0.24)	12 (0.16)	26 (0.42)	11 (0.13)	24 (0.32)	27 (0.46)	21 (0.30)
E	54 (0.88)	45 (0.76)	61 (0.84)	36 (0.58)	76 (0.87)	52 (0.68)	32 (0.54)	50 (0.70)
Total	61	59	73	62	87	76	59	71

I = intra-experimental intrusions; E = extra-experimental intrusions. The raw scores converted to proportions are shown in brackets.

Recall protocols were scored for the number of items per category recalled. These are shown in Table III, Row C. They indicate that there are higher scores in the HA and homogeneous conditions. Additionally more category labels are

recalled (Table III, Row D) in the 3 × 6 homogeneous HA and LA conditions than in the heterogeneous equivalents and this also applies to the LA 6 × 3 lists. The HA 6 × 3 lists do not follow this pattern but this difference can be located in an abnormally low level of performance on one of the two HA homogeneous 6 × 3 lists. It is interesting to note that in the percentage of recalled clusters preceded by their cue word (Table III, Row F) the homogeneous list score is greater, in every case, than the comparable heterogeneous score. If direct inter-item associations were the basis of the recall process and homogeneity led to no more than an equalization of the mean strength of the inter-item associative network, this effect would not be expected. The primacy of the core word recall suggests that it plays some part in the organization of the subsequent recall and it might well be argued that the presentation of the core word in the first position increases the probability that it will be given as the implicit associative coding response to the later words in the presented group where this is a homogeneous one. Thus these data would support a coding response process of some kind as being involved in the reception and storage of the material but seems incompatible with a view based solely on direct associative processes.

Discussion

(a) The finding that cues have no facilitating effect on the recall of the 3 × 6 lists but have a small amount of facilitation with 6 × 3 lists agrees with those of Dallett (1964), Tulving and Pearlstone (1966) and Weist (1970), that cues facilitate recall only where the number of categories is sufficient (greater than four) to allow forgetting of some of the category names to occur. It also accords with findings that the provision of cues at both learning and recall facilitates performance but that cue presentation at recall only does not (Tulving and Osler, 1968). The greater effect of cue provision in the HA and homogeneous conditions suggests that it is important that the provided cue is the one that is utilized by the subject. This correlation of effective and nominal cues is more likely in high-association and associatively homogeneous lists.

The smaller effect of cueing in heterogeneous lists is probably due to the ambiguous relationship of the core word to the rest of the set, for this was not always obvious and may have led to time being spent searching for the relevance of the cue to the succeeding words. Clearly the cues used here are different from the type used by Tulving and others, especially in the heterogeneous lists, in that they are not immediately obvious as focal points around which the rest of the category can be organized.

(b) The recall scores in the different experimental conditions produced by the non-cued recall subjects (with the cue word scores re-instated) showed the same tendencies as were reported in analysis 1. HA and homogeneous lists were better recalled than LA and heterogeneous lists. Additionally, recall was better in the 3 × 6 lists than in the 6 × 3 lists with the former showing a more efficient recall in the middle presentation positions than the latter. This suggests that ease of category identification or of an associative mediator may have been the dominant factor in any shift from the more usual finding in free recall studies of a higher failure

rate in the recall of items in the central presentation positions, and may have been particularly the case with the homogeneous lists where better recall in the centrally presented positions was most marked.   Further support for this view comes from the category recall data, i.e., the number of categories from which at least one item was recalled.   In the 3×6 lists this shows some facilitation in the HA homogeneous condition (see Table III, Row B).   This fails to occur in the 6×3 HA homogeneous lists but this is probably due to the abnormally poor performance (mentioned earlier) on one of the two lists used (86 categories recalled compared with 116 in the other list).   Similarly, recall of the category label showed an even clearer facilitation in the homogeneous condition and the incidence with which recall of a category was preceded by the appropriate core word was greater in the homogeneous lists than in the heterogeneous lists in every case.   Category label recall is in fact the best single predictor of total recall (see Table III, Row D).   For example, within-group inter-item association scores (Table III, Row H) would predict the same total recall scores for the HA heterogeneous and LA homogeneous conditions in both 3×6 and 6×3 list structures; category label recall scores are also the same for the HA heterogeneous and LA homogeneous conditions in 3×6 lists (where total recall scores are similar), but are higher for HA heterogeneous than for LA homogeneous conditions in 6×3 lists (where HA heterogeneous total recall is higher than LA homogeneous total recall).

Error data (Table IV) show related trends.   These were classified as either intra-experimental errors, i.e. items which had occurred in earlier presented lists, or extra-experimental errors, items which had not.   The latter were often semantically related to other items in the recall protocols, with the proportions of errors so classified ranging from 0·93 to 0·59.   The latter proportions and the proportions of total extra-experimental errors were greater in the homogeneous than the heterogeneous conditions in three out of the four conditions (HA 3×6; HA 6×3; LA 3×6).   The fourth instance in which this failed to occur was in the LA 6×3 condition, in which the difficulty of classifying the presented items into clearly definable semantic sets would have been greatest.   Classification of items into distinguishable sets is likely to reduce within-list confusion and lead to fewer intra-experimental intrusions.   If incomplete recall occurs, the availability of a "set" label seems likely to increase the production of semantically related intrusions.

These findings are of clear relevance to the views expressed in the introduction about coding processes.   If the subject uses only direct inter-item associations within a set as the basis for his recall processes, the stronger intra-category associative connections in the homogeneous lists should facilitate the recall of items within such a set as compared to the heterogeneous lists.   However, the order of recall should be less consistent in the homogeneous lists due to the strong inter-word associative links.   Because this fails to occur; because the best category label may be the core word in homogeneous lists, because this label frequently precedes the recall of the other items in the set in the homogeneous condition, and because category label recall is the best single predictor of recall efficiency, the importance of a coding process which may be reducible to the provision of a common implicit associative mediator or to a category label, rather than to the use of direct inter-

word associations, is suggested by the data. The differences in the error data between the homogeneous and heterogeneous conditions may also be interpreted as indicating that homogeneous lists may be coded more easily and that this is the basis on which items are produced in recall, so leading to more extra-experimental errors, which can be classified as members of a semantically related set of items.

These proposals about the facilitatory effect of labelling and the limitations on the storage capacity for coding labels can be further supported by data from the cueing section of the experiment. Cueing adds more to the conditions in which labels are likely to have been properly formulated (homogeneous and HA conditions) and may have been forgotten (6 $\times$ 3 lists).

References

BOUSFIELD, A. K. and BOUSFIELD, W. A. (1966). Measurement of clustering and of sequential constancies in repeated free recall. *Psychological Reports*, **19**, 935–42.

BOWER, G. H., CLARK, M. C., LESGOLD, A. M. and WINZENZ, D. (1969). Hierarchical retrieval schemes in recall of categorized word lists. *Journal of Verbal Learning and Verbal Behavior*, **8**, 323–43.

COFER, C. N. (1965). On some factors in the organisational characteristics of free recall. *American Psychologist*, **20**, 261–72.

COFER, C. N. (1966). Some evidence for coding processes derived from clustering in free recall. *Journal of Verbal Learning and Verbal Behavior*, **5**, 188–92.

COHEN, B. H. (1966). Some or none characteristics of coding behaviour. *Journal of Verbal Learning and Verbal Behavior*, **5**, 182–7.

DALLETT, K. M. (1964). Numbers of categories and category information in free recall. *Journal of Experimental Psychology*, **68**, 1–12.

DALRYMPLE-ALFORD, E. C. (1970). The measurement of clustering in free recall. *Psychological Bulletin*, **74**, 32–4.

DEESE, J. (1959). Influence of inter-item associative strength upon immediate free recall. *Psychological Reports*, **5**, 305–12.

JENKINS, J. J. and RUSSELL, W. A. (1952). Associative clustering during recall. *Journal of Abnormal and Social Psychology*, **47**, 818–21.

KIRK, R. E. (1968). *Experimental Design Procedures for the Behavioral Sciences*. Belmont, California: Wadsworth.

MATTHEWS, W. A. (1966). Continued word associations and free recall. *Quarterly Journal of Experimental Psychology*, **17**, 31–8.

MATTHEWS, W. A. (1969). Norms from continued controlled association. *Bulletin of the British Psychological Society*, **22**, 193–5.

SEGAL, E. M. (1969). Hierarchical structure in free recall. *Journal of Experimental Psychology*, **80**, 59–83.

TULVING, E. and PEARLSTONE, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, **5**, 381–91.

TULVING, E. and OSLER, S. (1968). Effectiveness of retrieval cues in memory for words. *Journal of Experimental Psychology*, **77**, 593–60.

WEIST, R. M. (1970). Optimal versus non-optimal conditions for retrieval. *Journal of Verbal Learning and Verbal Behavior*, **9**, 311–6.

WOOD, C. (1969). Retrieval cues and the accessibility of higher order memory units in multi-trial free recall. *Journal of Verbal Learning and Verbal Behavior*, **8**, 782–9.

THE CAPACITY FOR RAPID SHIFTS IN LEVEL OF MENTAL CONCENTRATION

GERALD S. BLUM AND MARCIA L. PORTER

Department of Psychology, University of California, Santa Barbara

This investigation continues the exploration of a largely neglected topic in the field of attention, levels of mental concentration, by ascertaining under optimal conditions the capacity for shifting quickly from one level to another. Previous research in our laboratory had applied a set of five post-hypnotically cued levels of concentration or "cognitive arousal", covering a very wide range, to a variety of perceptual and cognitive tasks. In the present study two highly trained undergraduate hypnotic subjects, one male and one female, were given progressively shorter time durations in which to respond to the post-hypnotic cues before viewing tachistoscopically flashed consonants under cue influence. The results indicate a striking capacity to shift degrees of cognitive arousal very rapidly to one extreme or the other, on the order of 100 ms or less.

Introduction

Psychologists, despite their rekindled interest in the topic of attention, have been singularly inattentive in recent years to the dimension of mental concentration. The degrees of attention capable of being invested in a particular stimulus have largely been ignored in the quest for principles governing selective attention in situations with competing inputs. Titchener (1908) emphasized the sensory attribute of *clearness*, later renamed "vividness" and then "attensity", but current reviewers do little more than list mental concentration as one among many categories of attention and comment upon the paucity of work done (e.g. Moray, 1969).

An exception is a series of studies in our laboratory, where for a number of years we have been pursuing the concept of *cognitive arousal*. Subjects have been trained under hypnosis to respond to five post-hypnotic cues with degrees of attentiveness ranging from a peak of concentration and alertness all the way down to a level approaching stupor. These degrees of mental concentration have now been applied in a wide variety of settings which point to the construct validity of the prior hypnotic programming and rule out such confounding variables as the operation of demand characteristics. For example, degrees of cognitive arousal have consistently yielded ordered relationships to: response latency in a perceptual discrimination task (Blum, Geiwitz and Stewart, 1967); degrees of boredom (Geiwitz, 1966); strength of cognitive reverberation of strings of consonants (Blum, Geiwitz and Hauenstein, 1967; Blum, Hauenstein and Graef, 1968); degrees of organization of thought processes (Blum, Graef and Hauenstein, 1968); and performance on the Stroop Color-Word Test (Blum and Graef, 1971).

In the above studies subjects were typically given approximately three seconds to reach the appropriate degree of arousal following presentation of a post-hypnotic cue. This period of time proved more than ample to reach either extreme, judging by both subjective reports and observed experimental effects. The question remains whether the alertness shifts can be accomplished in an even shorter time interval. The present study is designed to push the capacity for shifting concentration to its limits under optimal conditions, so that we can establish minimum required times to move up and down the scale of alertness.

Method

Subjects

Two subjects in our laboratory were ideally suited to serve in a test of the capacity for switching levels of concentration in the shortest period of time. The first, a 20-year-old female undergraduate (F), had participated as a paid subject in a variety of hypnosis experiments for over two years with two or three sessions a week. The second, a 20-year-old male undergraduate (M), had been employed similarly for over one year. Both had initially scored the maximum of 12 on the Stanford Hypnotic Susceptibility Scale, Form A, and had demonstrated exceptional proficiency on a series of advanced hypnotic tasks. They were already very experienced in responding to five post-hypnotic cognitive arousal cues. In the waking state, while amnesic for the prior hypnotic programming, they became exceptionally alert upon visual presentation of the cue "+AA" and almost stuporous in response to "−AA". A medium state of arousal was cued by "O", halfway between O and +AA was labelled "+A", and halfway between O and −AA was labelled "−A". Details of this training procedure, along with methods for validation and for ruling out simulation and operation of demand characteristics, are available elsewhere (Blum, 1972).

Procedure

The following task was set up to vary the time period of cue action and thereby ascertain the minimum duration at which that influence is still observable. A trial contained five cycles of tachistoscopic exposures followed by a report period. Each cycle consisted of an arousal cue shown for a specified duration, followed by the flash of a consonant, and then by a two-digit number in response to which the subject began counting backward by threes, as fast as possible without making a mistake, until a second arousal cue appeared to start the second cycle. The counting interval between appearance of the number and the next cue was 7 s. After the five viewing cycles, which paired the 5 arousal cues with 5 different consonants, were completed, a blank flash accompanied by the first of a series of clicks at 1-s intervals signalled the subject to begin his report. The latter consisted of the verbalization of whatever consonant came into mind at each of the next five successive clicks. The fate in the report of each cue-paired stimulus consonant served as the basis for determining whether or not the cue had sufficient time to act.

Subjects were told previously under hypnosis to respond immediately to each mental arousal cue and to maintain that level of concentration until the following consonant disappeared (i.e. the counting number appeared). In other words, the appropriate level of arousal was to be held throughout the presentation of the consonant but no longer. Amnesia was instructed for the prior hypnotic programming and no mention of the cues was made in the waking task directions. In order to facilitate the independence of trials, subjects were also instructed to be amnesic for what had transpired as soon as each set of five cycles was completed.

Apparatus and stimulus materials

The subject was seated at an Iconix 3-field tachistoscope in a semi-darkened room. Before each trial two experimenters preloaded Field A with five cue cards (−AA, −A, O, +A, +AA drawn 1 in. (2·54 cm) high in black ink) stacked in one of eight random orders;

Field B with five different consonants ($\frac{3}{16}$ in. (0·48 cm) high, stamped from a printing set in black ink) arranged in a predetermined order; and Field C with five randomized cards each containing a different two-digit number between 50 and 99 (same format as consonants). The order of consonants preloaded in Field B was adapted from an already available pool of 432 hexagrams constructed according to the following criteria: no consonant repeated within a hexagram; no consonant repeated in two consecutive hexagrams; no sequences with apparent meaning; and omission of V and N (originally because of their acoustic confusability with B and M in an auditory task). Q was also omitted in the present investigation because of its possible association to "cue", leaving a total number of 17 consonants. Viewing distance to a 4 × 4 in. (10·16 × 10·16 cm) exposure area when the subject placed his head in the hood was 40 in. (1·016 m) and illumination in the three fields was 35, 71, and 30 ft lamberts respectively (1 ft lambert = 3·42626 cd/m²). Rise and fall time of the fluorescent lamps was less than 500 μs. Fastened below the hood, just beneath the subject's chin, was ostensibly a microphone, which in actuality triggered a voice relay in order to time the latency of the subject's responses without his knowledge. Lag time between vocalization and triggering was less than 1 ms. Other microphones located about the room did serve to record the content of the subject's vocalizations.

Subjects came to the laboratory twice a week and a session consisted of 40 trials (a trial included five cycles of cue, consonant, and counting, plus the report) at a given cue duration. The subject was given a 5-min rest after every 10 trials in addition to a 30-sec rest between trials. One experimenter manually changed the five cue cards in Field A during the course of a trial, recorded the latency in hundredths of a second of the subject's first counting response in each cycle, and also wrote down the five consonants which the subject reported at the end; the other experimenter changed the five consonant cards in Field B and the five number cards in Field C as the trial progressed.

Experimental design

We have already mentioned some characteristics of the design. Every trial paired each of the five arousal cues with a consonant. What remains to be described is the manipulation of cue duration, the independent variable. In the first part of the experiment the consonant was always flashed for 0·5 s. The cue exposure (which is not the same as cue duration) was initially set at 2·5 s and progressively shortened. Thus cue duration, i.e. cue exposure plus consonant exposure, began at 3·0 s and decreased thereafter.

Later in the experiment, after it had been determined that a cue duration of 0·7 s was still sufficiently long to yield systematic effects of the arousal cues, it was obviously necessary to begin shortening the consonant exposure as well. After the combination of 0·1 cue exposure and 0·3 consonant exposure again produced positive results, cue exposure was then set constant at its minimum assured visibility of 0·05 s and consonant exposure was progressively shortened until the effects vanished.

Results

Arousal Cue Effects

Table I shows the number of times out of 40 trials that each subject reported consonants associated with each arousal cue at the progressively decreasing cue durations. The cues have an obvious effect for F down to the duration of 0·15 s and for M down to 0·1 s.

For F, the mean frequencies for all cue durations except 0·1 s are: —AA, 13; —A, 13; O, 19; +A, 25; +AA, 28. These differences are statistically significant ($F = 39·21$, $P < 0·01$). Newman-Keuls comparisons of the means for the five arousal levels show —AA and —A different from O, +A, and +AA; +A and +AA different from O, —A, and —AA. In other words, there are three separable

TABLE I

Number of reported consonants in each of five arousal levels at varying cue durations for subjects F and M

Cue duration (s)				Arousal levels −AA	−A	O	+A	+AA
3·0	Cue	2·5	F	11	16	19	22	30
	Con	0·5	M	7	8	33	31	29
2·0	Cue	1·5	F	12	13	19	21	30
	Con	0·5	M	8	5	36	36	35
1·5	Cue	1·0	F	14	18	22	25	31
	Con	0·5	M	4	4	35	37	39
1·0	Cue	0·5	F	11	14	27	25	25
	Con	0·5	M	6	7	36	36	39
0·7	Cue	0·2	F	16	13	22	27	29
	Con	0·5	M	7	6	38	39	38
0·4	Cue	0·1	F	12	7	12	35	31
	Con	0·3	M	10	5	37	36	39
0·3	Cue	0·05	F	17	16	17	24	30
	Con	0·25	M	5	9	38	38	39
0·25	Cue	0·05	F	13	8	23	24	27
	Con	0·20	M	6	2	40	39	38
0·2	Cue	0·05	F	9	13	15	25	26
	Con	0·15	M	13	4	38	38	37
0·15	Cue	0·05	F	13	14	16	25	23
	Con	0·10	M	5	4	33	35	33
0·1	Cue	0·05	F	12	10	13	15	9
	Con	0·05	M	10	6	29	30	29
0·075	Cue	0·05	F	—	—	—	—	—
	Con	0·025	M	8	11	11	12	12

Note: Maximum number possible in each cell is 40.

levels of influence: the minus cues, O, and the plus cues. At 0·1 s there are no significant differences.

For M, the results are similar. Excluding the 0·075 duration, the mean frequencies are: −AA, 7; −A, 5; 0, 36; +A, 36; +AA, 36. The differences among cues are significant ($F = 525 \cdot 16$, $P < 0 \cdot 01$), but only two levels of effect are apparent: the minus cues versus the others (O, +A, +AA). At 0·075 s the effect of the cues no longer obtains. A closer look at M's responses suggests that 0, +A, and +AA may indeed have acted differentially in his case, but that the variable of report frequency was not sufficiently sensitive to detect these differences. Analysis of mean *order* of occurrence within the reports of five consonants indicated that the +AA consonant occurred significantly earlier than the O consonant, with the +A consonant falling in between.

Latency data

Another measure was the latency of each subject's first backward counting response to the two-digit number. Recorded by means of a voice-operated relay,

these data were gathered to provide an empirical check on actual cue duration. By instruction, the cue was supposed to go off with the simultaneous offset of the consonant and onset of the number, but there is the possibility that the arousal state lingered beyond that time. If so, the speed of response would be faster for the plus cues, slower for the minus cues. This direct effect of the cues on latency has been shown in prior research (e.g. Blum, Geiwitz and Stewart, 1967).

For F, the mean latencies in seconds for the various cues are: —AA, 1·42; —A, 1·41; O, 1·49; +A, 1·52; +AA, 1·53. For M: 1·36, 1·33, 1·89, 1·88, 1·73, respectively. These means are significantly different for both subjects ($F =$ 4·62 and 44·34 respectively). The speed of response for the minus cues is clearly greater than for the plus cues, the exact opposite of a prediction based on the lingering influence of the cues.

Inquiry reports by subjects

Comments made by the subjects in waking and hypnotic inquiries upon conclusion of the experiment are of some interest. F reported that she was able to see the consonants in the 0·15 condition (exposed for 0·10 s) but was unable to do so in the 0·1 condition (exposed for 0·05 s). The cues themselves, printed in larger, thicker letters, were visible throughout. M, on the other hand, said that he could still see some of the consonants in the 0·1 condition but not in the 0·075 condition (exposed for only 0·025 s). These waking comments suggest that the constraint of visual acuity precluded cue effects at 0·1 for F and at 0·075 for M, the cue durations where significant differences were no longer found.

Under hypnosis, F stated that, with the plus cues, the consonants made a "strong impression in my mind"; with the minus cues they were "garbled, hardly distinct at all". M reported that the consonants were "especially sharp" under the influence of +AA and only a glance would suffice for them to register clearly. At the lowest level of concentration, —AA, his mind became so drowsy that he had to force himself to look at the centre of the screen in order to try to see the upcoming consonant, which he knew he had to do.

Simulation data

After the above patterns of results, including report frequencies and response latencies, were known for the two subjects, it became of ancillary interest to ascertain whether they could simulate their obtained data on a consciously deliberate basis. Accordingly F, in a series of sessions with only waking instructions, was strongly encouraged to reproduce, by whatever means possible, a specified distribution of reported consonants over each set of 40 trials. The task followed the same format as before except that +AA was replaced by the number "1" (whose associated consonants were to appear in the report on approximately 30 out of the 40 trials); +A replaced by "2" (its consonants to be reported on 27 trials); O by "3" (20 reports); and —A and —AA by two "4"s (13 reports each). Four sessions were conducted, ranging from "cue" durations of 3·0 s through 0·2 s. Overall, F was very successful in reproducing ordered distributions of consonant reports

which did not depart markedly from her designated numerical objectives. However, it became clear that she had not employed those same techniques in the experiment proper: the counting latencies no longer showed any difference among conditions (averaged over sessions, all four cue numbers fell between 1·33 and 1·37 s); and the relative frequencies of cue-associated consonants reported within a single trial often differed radically from the original data.

M's simulation sessions contained the instruction to omit completely from the report on each trial those consonants associated with the two number "6"s, and to report those associated with "1", "2", or "3" in the order of their exposure (a consistent pattern which had emerged for O, +A, and +AA near the end of the experimental series). Time durations of the "cues" ranged from 3·0 s down to 0·1. At the slowest speed he had fairly good success in carrying out the instructions but his performance deteriorated as the exposures grew shorter, so that by 0·1 the report frequencies no longer bore any similarity to his original data. The response latencies also did not conform to the experimental results.

It is noteworthy that both subjects stated afterward that they tried to rehearse at every opportunity during the simulation task, whereas they never rehearsed in the experimental series. The fact that the two conditions turned out to have basic differences is consistent with their observations.

Discussion

One issue which must be clarified concerns the possibility that, in the experiment proper, the subjects deliberately engaged in varying amounts of rehearsal during the counting backward task, so as to produce seemingly systematic effects of the arousal cues. We have just noted that the simulation data do not support such a contention. Also, both subjects, in waking and hypnotic inquiries after the main experiment, asserted that they followed instructions and never rehearsed the consonants at any point prior to reporting. There is no reason to doubt their honesty on the basis of their past histories in our laboratory. Honest reporting, whether or not the experienced phenomena conform to hypnotic or post-hypnotic instructions given by the experimenter, is stressed in both waking and hypnotic states at the very outset of their participation. Subsequent training sessions include a variety of checks on simulation (see Blum and Graef, 1971) and in those sessions subjects frequently report failures in contradiction to the instructional demands placed upon them. Similarly, Bowers and Gilmore (1969), from a complex experimental design testing the credibility of reports of hypnotic hallucinations, conclude that their data suggest that hypnotic subjects are honest and reports are credible across a variety of situations.

If we grant that conscious rehearsal was not attempted, how, then, are we to account for the latency findings which reveal acceleration of the onset of counting for the two low-arousal cues? The latency measure was included originally as an empirical check for any possible delay in cue offset, which can in fact be ruled out by the results. The occurrence of shorter counting latencies following the minus cues is of course compatible with the notion of conscious rehearsal, which would not be undertaken for −A and −AA if the subject were inclined to minimize the

likelihood of their appearance in the report. An alternative explanation, equally compatible with the data and at the same time congruent with the assumption of subject honesty, can be given in terms of weaker persisting interference of the recently exposed consonant with the process of counting. If lowered arousal weakens registration of its associated consonant, the latter is less likely to be reverberating when counting has to begin, and latencies are consequently shorter. To check on this interpretation, the prediction was made that, within the −AA level, those associated consonants which do make their way into the report should interfere more with the start of counting than those consonants knocked out completely. Consistent with this view, mean latencies of −AA consonants subsequently present in the report for subject F were 1·56 s compared to 1·46 s for −AA consonants not reported ($t = 2·37$, $P < 0·01$); for M the respective figures were 1·54 and 1·39 ($t = 1·86$, $P < 0·05$). In essence, this alternative explanation invokes a concept of unconscious, autonomous reverberation, which has been strongly supported in earlier studies in our laboratory (Blum, Geiwitz and Hauenstein, 1967; Blum, Hauenstein and Graef, 1968).

It seems reasonable to conclude, therefore, that our subjects were indeed capable of very rapid shifts in level of concentration, on the order of 100 ms or less. For F the arousal cues were still producing differential effects at a duration of 150 ms; for M at 100 ms. Very likely these limits were not exceeded because F's vision was not adequate to discriminate consonants flashed for only 50 ms, and M could not recognize them at 25 ms. Apart from these visual constraints, it can be assumed that the obtained shift durations represent overestimates since the prerequisite time to identify the cue card itself should be subtracted.

At the outset we stressed that our experiment was designed to test, under optimal conditions, the capacity for shifting concentration. The step-by-step decrease in exposure durations intentionally capitalized on the effects of practice. Certainly the use of hypnotic programming is not a conventional mode for altering cognitive arousal, nor can our two highly susceptible hypnotic subjects, trained and practised over long periods of time, be considered typical. Nevertheless, there is no reason to assume that the basic mechanisms involved in such shifting are unique in their case. Rather we can view the hypnotic manipulation as sharpening, refining, and controlling more precisely a capacity which everyone does possess in varying degree. In that sense the experimental test is optimal for determining minimum time required to move up and down the scale of alertness.

On the other hand, it may be that momentary shifts in concentration do occur normally and repeatedly in everyday life under fine control of internal messages of which we are unaware. Perhaps it is not too far-fetched to conceive of a sudden and fleeting loss of concentration as a defensive measure in response to a subtle warning signal of anxiety, depriving the associated content of an opportunity to achieve conscious awareness (repression?). Also, a momentary loss of concentration may very well be implicated in the execution of post-hypnotic amnesia.

References

BLUM, G. S. (1972). Hypnotic programming in psychological experiments. In FROMM, E. and SHOR, R. (Eds), *Hypnosis: Research Developments and Perspectives.* Chicago: Aldine-Atherton.

BLUM, G. S., GEIWITZ, P. J. and HAUENSTEIN, L. S. (1967). Principles of cognitive reverberation. *Behavioral Science,* 12, 275–88.

BLUM, G. S., GEIWITZ, P. J. and STEWART, C. G. (1967). Cognitive arousal: the evolution of a model. *Journal of Personality and Social Psychology,* 5, 138–51.

BLUM, G. S. and GRAEF, J. R. (1971). The detection over time of subjects simulating hypnosis. *International Journal of Clinical and Experimental Hypnosis,* 19, 211–24.

BLUM, G. S., GRAEF, J. R. and HAUENSTEIN, L. S. (1968). Effects of interference and cognitive arousal upon the processing of organized thought. *Journal of Abnormal Psychology,* 73, 610–14.

BLUM, G. S., HAUENSTEIN, L. S. and GRAEF, J. R. (1968). Studies in cognitive reverberation: replications and extensions. *Behavioral Science,* 13, 171–7.

BOWERS, K. S. and GILMORE, J. B. (1969). Subjective report and credibility: an inquiry involving hypnotic hallucinations. *Journal of Abnormal Psychology,* 74, 443–51.

GEIWITZ, P. J. (1966). Structure of boredom. *Journal of Personality and Social Psychology,* 3, 592–600.

MORAY, N. (1969). *Attention: Selective Processes in Vision and Hearing.* London: Hutchinson Educational Ltd.

TITCHENER, E. B. (1908). *Lectures on the Elementary Psychology of Feeling and Attention.* New York: Macmillan.

VISUAL AND VERBAL CODES: SPATIAL INFORMATION SURVIVES THE ICON

LESLIE HENDERSON*

Department of Psychology, University of Waterloo, Ontario

Three experiments are reported to demonstrate partial independence of identity and spatial position information concerning visually presented symbols. Experiment I shows that performance on these forms of information improves at different rates as a function of exposure duration. Experiment II shows that performance on one can be traded against the other. Experiment III demonstrates partial statistical independence of item and position responses and shows that increases in the duration and delay of the probe facilitate performance. Some implications of these experiments for theories involving mandatory verbal encoding of visual symbol-arrays are discussed. A model is proposed comprising visual and name stores with different acquisition rates and capacities. Both are indexed by identity but the visual code can also be interrogated by spatial cues.

Introduction

William James, in the section of the Principles devoted to Primary Memory, treats two sorts of visual memory. The first is a pre-attentive store or sensory register which we shall call "iconic memory", following Neisser (1967). The second form of visual storage is implied in James' discussion of certain observations of Fechner's (James, 1890, p. 645). This visual code is subsequent to attention and lacks the sensory image qualities of iconic memory.

Several modern theories of visual information processing (Sperling, 1970; Atkinson and Shiffrin, 1969) have retained the notion of an iconic memory but assume that symbolic material is transferred to verbal memory by implicit speech in order to survive the icon.

One difficulty with this view is the success with which position information is reported after iconic decay (e.g. Sperling, 1960). In order to recode such information into verbal STM either to each item must be added descriptions of its location or a systematic scan path must encode position information as serial order.

The first alternative seems highly implausible. In terms of Sperling's (1970) model, the verbal descriptions of position would require a revisal of the estimated scan rate and recognition buffer capacity. More obviously, processing such descriptions would make overwhelming demands on the limited capacity rehearser and auditory store. From the foregoing it is evident that relaxing the demand for positional accuracy should effect a marked improvement in performance. No such evidence exists; nor can the subject report any such coding.

If a fixed scan path encoded position then it should produce certain systematic effects on performance as a function of spatial location. It is difficult to see why

*Now at University of Guelph, Guelph, Ontario, Canada.

these functions should be as idiosyncratic as those of Sperling (1967). The Scanner must follow some predetermined vectors and therefore would have to enter blank data from certain locations in the parallel-processed icon, with implications for a limited capacity STM.

A paradigm for independent measurement of item (I) and position (P) information has been devised by Cumming and Coltheart (1969) who conclude that I and P information are stored as a unit and forgotten or retained together. The authors' note, however, that their results are also compatible with the hypothesis that correct recall of I entails recall of P but not the converse. Furthermore average accuracy in this study was 1·2 digits out of a six item array. The authors note that such data may not be generalizable to higher performance levels.

In view of the equivocal nature of Cumming & Coltheart's (1969) results and the theoretical difficulties entailed by acceptance of their assumption that I and P information are stored as a unit in Sperling's (1967) auditory store the following experiments were performed using the same paradigm.

Stimulus Material and General Procedure

One hundred and eighty stimulus- and response-array pairs were constructed from black Letraset and Chart-Pak tape transferred to white cards. Both the stimulus- and response-arrays contained a 3 × 3 cell matrix in the form of a "noughts and crosses" grid. In the stimulus-array five cells were filled at random with digits randomly sampled from the set 0–9 without replacement. The response-array for a given stimulus-array contained two cells which were filled, one with a £ symbol and one with a $. To the left of this array were two digits one above the other. Exactly one of these digits had occurred in the stimulus-array and exactly one of the symbol-designated locations had been filled. In fact that location had been filled by that digit. The subject's task was to make a two-bit forced choice response indicating which digit had occurred in the stimulus and in which location. Which symbol was correct on a given trial and which digit (upper or lower) was varied randomly. Over the entire stimulus set, each location in the matrix was probed an equal number of times.

The stimulus-array was displayed in field 1 of a Gerbrands 2-field tachistoscope. At the viewing distance of 560 mm the matrix subtended 9·1°, each cell being a 3·0° square. The digits had an angular height of 1·0° and were centred in the cell. Background luminance was about 13·5 cd/m².

Field 2 contained a visual noise display constructed to serve both as a fixation indicator and a mask. It contained a grid identical to that of the stimulus-array, aligned to yield superimposition with the stimulus. The cells of this grid contained visual noise in the form of randomly disposed fragments of the digits used in stimulus-arrays.

The response-arrays were beside the tachistoscope so that the subject could readily turn from the apparatus to consult the appropriate response-array. The two responses were spoken by the subject and recorded by the experimenter. After an inter-trial interval of about 8 s the ready signal heralded a new trial.

Experiment I

Procedure

The general procedure and materials were as described. The sequence of events for a trial was as follows. A small transilluminated point in field 2 served as fixation marker and was present throughout as a dim red dot. When the subject pressed the initiate button the blank display was replaced by the stimulus field for 400, 600 or 1000 ms according to the condition. The stimulus offset was coincident with the onset of the mask field which lasted

for 1 s and then the blank display returned. The offset of the mask field signalled to the subject that he should consult the response-array.

The exposure duration conditions were presented in blocks of 60 trials. The order of conditions was counter-balanced so that each condition occurred three times in each serial position across the nine subjects. All three blocks were run in a single session of about 1h. At the beginning of the session the subject was given 20 practice trials at the shortest duration. Before each block he was given a few examples of the duration to be used.

Subjects

The subjects were nine undergraduates at the University of Guelph, five females and four males.

Results

All analyses were performed on raw scores. References are made in the text to scores corrected for chance guessing where the absolute level of performance is of interest. The correction is for one bit forced-choice responses. The data are only examined with respect to the effect of exposure duration on item and position performance. That information is presented in Table I, which shows item and positions scores out of 60 for each of the stimulus durations. Item performance seems to improve more than position performance as exposure duration is increased. This proposition was tested by subjecting the I-P scores to a Friedman two-way analysis of variance by ranks for the 9 subjects $\times$ 3 durations. It was found significant with $\chi r^2 = 6.4$, $P < 0.05$.

TABLE I

Mean item and position scores (out of 60) for three exposure durations of the stimulus. Superiority of item performance is expressed as Item minus Position score. Scores are not corrected for guessing. Item = I; Position = P

400 ms			600 ms			1000 ms		
I	P	I-P	I	P	I-P	I	P	I-P
50·7	50·3	0·4	54·2	50·9	3·3	56·1	52·1	4·0

Inspection of the group means shows that accuracy, corrected for guessing, rises from 68% to 74% for position and from 69% to 87% for items. For some subjects a ceiling effect is likely to have attenuated effects of increasing duration.

It appears, then, as if increasing exposure duration affects item and position performance differentially.

Experiment II

Procedure

The general procedure and materials were as before. In this experiment a single-channel tachistoscope timer was used and the noise field served as a fixation indicator and mask. The subject fixated in the centre of the mask and when a trial was initiated the mask was replaced by the stimulus array for 250 ms at the expiry of which the mask returned. The subject then consulted the response-array. A shorter duration than in Experiment I was employed to bring performance into the central portion of the psychometric function.

There were two instructional conditions. Subjects were either instructed to adopt a processing priority in favour of item or of position information. They were to devote all their attention to the instructed information, attempting to do as well as possible on it and

contenting themselves with doing as well on the other form of information as this priority permitted. They were also instructed to report the high priority information first. There were 90 trials in a single block for each condition and conditions were run in separate sections.

Subjects

Twelve University of Guelph undergraduates served as subjects, six of each sex.

Results

The data are summarized in Table II. Overall position information is transmitted more successfully than item ($F = 9 \cdot 4, P < 0 \cdot 02$). Even with item-priority instructions, there is a tendency toward better performance with position information.

TABLE II

Item and position information transmitted as a function of instructions.   Mean scores out of 90

| Item | | Position | | (Instructions) |
I	P	I	P	(Information)
69·2	71·5	66·3	75·6	

The interaction effect is significant ($F = 9 \cdot 9, P < 0 \cdot 01$). Subjects can reliably trade item for position processing. In absolute terms the trading limits are small.

A few subjects spontaneously reported that position-priority instructions were difficult to follow as it seemed more "natural" to give priority to item information in a task of this sort. This suggests that the trading limits might be slightly underestimated here and that the superiority of position performance might also be so.

Discussion

Table III summarizes the results of both experiments. The group mean data have been transformed to per cent accuracy corrected for guessing. The experiments differ procedurally but some estimate of the effects of exposure duration

TABLE III

Group mean per cent correct performance for Experiments I and II.   (Scores are corrected for chance guessing)

| Instructions | Experiment II | | Experiment I | | | | | | Stimulus duration (ms) |
| | 250 | | 400 | | 600 | | 1000 | | |
	I	P	I	P	I	P	I	P	Type of information
Item priority	54	59							
None			69	68	80	70	87	74	
Position priority	47	68							
(Collapsing conditions in Experiment II)	51	64	69	68	80	70	87	74	

across experiments can be surmised. There is a clear suggestion that the inter-action between type of information and duration shown in Experiment I might be extrapolated to a cross-over of P and I performance at shorter durations. One explanation for this might be that position information is acquired by a rapid acquisition visual code with a limited capacity. This code might be an unstable medium for item information which relies more on a name code with larger capacity but slower acquisition rate. Such an account has been advanced by Scarborough (1971) in a very different experimental situation.

The results support the notion that item and position information are at least partially independent. While it must also be conceded that performance never differs greatly between I and P in any of the conditions explored, a limited dependency could be mediated by quite irrelevant constraints. If locations were to be scanned sequentially, for example, and item and position information encoded from the location into independent codes then performance levels would tend to be similar throughout the psychometric function. In fact position information would grow faster at first than item and then slower. This follows from the fact that an empty location carries position information but no item information. Consider, for example, the case of three locations sampled and two of them containing items. If an item or position probe refers to one of the occupied cells then the information pertinent to a correct response will have been acquired. If a position probe refers to the empty cell then the information for a correct response will also have been acquired since the alternative position probe must be correct even though its location has not been examined.

Such a model serves to illustrate that the experimental paradigm cannot in principle corroborate the contingent coding proposition. In this it is asym-metrical, for it can and apparently does, support the converse.

Experiment III

In Experiments I and II the response-array had been presented beside the tachistoscope and the subject had to make a head movement to consult it. Thus probe-delay and probe-duration were not controlled. In fact, the subjects appeared to study the probe for some time. The locations on the response-array had been specified by monetary symbols. These are not readily confused with digits. However, it is possible that the use of symbols imposes an unnecessary burden in processing the probe. The time taken to process probes has received little attention though it is quite evident that this is an important factor in iconic memory. It is clear that negative probe delay improves performance in iconic memory even when the prior probe cannot direct an eye movement (Averbach and Sperling, 1961; Dick, 1969).

The present experiment brings these temporal parameters under control. The delay and duration of the probe are manipulated. Probe locations are specified by coloured patches. This not only eliminates the use of symbols for position but minimizes the backward masking due to presenting the probe in the same location as the item.

Finally, both exposure duration and display size were reduced slightly to preclude any role of eye movements in stimulus processing.

Material

Forty-five stimuli were constructed in similar manner to those of the previous experiments. Each of the nine matrix locations contained the correct item and position on five occasions. The item and position specified by the "incorrect" probes were randomly chosen. The response-arrays consisted of the matrix grid, as before. The two probe digits were to the left of the grid, vertically disposed. Position probes consisted of patches of matt red or blue colour which almost filled the grid cell. These patches were fairly dark and were chosen to be of approximately equal reflectance.

The third array contained either a blank white card with an outline square or a square of the same size filled with visual noise. The noise consisted of a high contrast photograph of the microstructure of a sheet of coarse sandpaper which, when superimposed on the stimulus, rendered it illegible. These two displays define different experimental conditions and will be referred to as "blank" and "noise" respectively.

Procedure

The stimuli were delivered by means of a Scientific Prototype 3-field five channel tachisto-scope. At the viewing distance of 1250 mm the digits each subtended an angular height of 50'. A single cell of the grid subtended 2·0°. The fields were aligned so that, with all fields illuminated, the stimulus and response grids superimposed and the blank or noise square circumscribed the grid and was concentric with it. At the luminances used legibility of the digits with probes superimposed did not apparently differ from that of other digits.

A trial consisted of the following sequence. The subject was initially confronted with a square, either blank or containing visual noise. This indicated where the stimulus would fall and the subject was instructed to fixate the centre of the square. At a ready signal the subject depressed a hand held trigger which caused the stimulus-array to replace the square for 150 ms. The square, be it blank or noise, returned for a "probe delay" interval of either 50 or 500 ms and was then replaced by the probe for a duration of 500 or 2000 ms. Finally the square returned until a new trial was initiated.

The background luminance of noise, blank and stimulus fields was 68 cd/m²; that of the probe field was 34 cd/m². Contrast of the digits was approximately 10%. Luminance was monitored throughout the experiment with a Spectra photometer.

Noise or blank fixation/masking field was a between subject variable and subjects were assigned to either condition in alternating order of arrival at the laboratory.

The two levels each of probe delay and probe duration were combined factorially as within subject variables and the order of treatments was so counterbalanced that each combination of treatments occurred equally often in each serial position.

Each combination of treatments consisted of a block of 45 trials, one trial with each stimulus-response pair; the order of stimuli was randomized. Subjects performed four blocks equally divided between two sessions, each of about 1 h. Each session began with several practice trials.

Subjects

Twenty-four subjects were run in the experiment. Of the 12 males, six were allocated to each of the noise and blank conditions. The subjects were recruited from the Introductory Psychology subject pool of the University of Guelph.

Results

The frequency of correct responses for item and position information was calculated for each block of 45 trials. These scores were subjected to a four-way analysis of variance for 24 subjects. Of the four two-level factors, type of informa-tion, probe delay and probe duration varied within subjects and noise versus blank varied between subjects.

Rather surprisingly there was no suggestion of a main effect of noise as pre-exposure and mask field. There was, however, an interaction between noise and type of information ($P = 0.002$) such that position performance was impaired by noise. This observation is consistent with the effect obtained by Liss (1968).

The other main effects were significant. Position performance was better than item ($P = 0.05$). This effect, while fairly reliable, was small in magnitude, the overall average performance for position and item information corrected for guessing being 40·4% and 35·2% respectively.

Performance was better at the longer probe duration ($P = 0.02$) and with the longer probe delay ($P < 0.005$). None of the remaining interactions approached significance (P min > 0.2).

The data were examined with respect to performance as a function of the location of the correct responses. The frequency of correct responses was summed across subjects and trials for each of the nine matrix locations. This yielded a topographic description of performance for each condition. The patterns obtained had great ordinal similarity across conditions. The coefficient of concordance (Kendall) was computed for the nine cells ranked by performance for each of the 16 conditions. This yielded $\chi^2 = 56.1$; $df = 8$, $P < 0.0001$. A summary of this topographic data collapsed over conditions is presented in Table IV.

TABLE IV

Topographic analysis of performance. Per cent accuracy averaged over all conditions

70	67	64
74	77	69
67	63	63

To test for unilateral dependencies between item and position performance the hypotheses (i) that item wrong entails position guess (HPI), and conversely, (ii) position wrong entails item guess (HIP) were tested in the manner of Cumming and Coltheart (1969). The frequency of the contingencies I_wP_c, I_wP_w and I_cP_w was summed across subjects, where the subscripts w and c denote wrong and correct performance on the given information. These totals were 800, 551 and 677 respectively. χ^2 tests of I_wP_c and I_cP_w against I_wP_w both demonstrated ($P < 0.001$) that performance on each form of information is significantly better than chance when the other is wrong. This entails rejection of the hypothesis of complete contingency. The disagreement between these findings and those of Cumming and Coltheart may be accounted for by the extremely low level of their subjects' performance. Maximum information about independence is transmitted at the middle of the psychometric function. As performance approaches chance (or perfection) statistical dependencies are introduced by that fact. (Cumming and Coltheart's subject JS who shows clearly the best performance also exhibits evidence of independence. On the other hand two of their subjects actually do worse than chance on one form of information.)

Discussion

The failure of the sandpaper-pattern visual noise to mask more effectively than the blank white-field is probably attributable both to the luminance summation masking provided by the relatively bright blank field and the relative lack of similarity between the components of the noise and the stimulus array. Such similarity has been related to masking efficacy by Liss (1968) and Henderson, Coles, Manheim, Muirhead and Psutka (1971) who suggest that masks with similar components to the stimulus have their effect by competition for feature analysers.

The superiority of performance with the longer probe delay may be due to a more complex form of such competition for the central processor. The arrival of the probe appears to be frustrating. If the subject attends to the probe, which must be processed, he thereby risks losing his grip upon the information maintained in memory. Sometimes a subject will report such success in attending to the memory code that the probe will come and go unattended leaving him with information that he does not know how to report. Delaying the probe seems to give the subject time to prepare for the demands it will place on him. This may give him time to transform the memory into a readily interrogated form, or it may simply allow some cognitive operations to take place which increase the resistance of the code to interference from processing the probe.

The longer probe duration may have its beneficial effect in part for similar reasons. The 2-s probe is less peremptory in its demand for attention. However, the lack of interaction between probe delay and duration suggests that the principal effect of probe duration is quite simply that it allows more time for the probe to be processed without having to allocate memory capacity to the probe itself. The fact that such processing may extend beyond 0·5 s is consistent with Dick's (1969) finding that performance on an iconic memory task improved when probe delay was increased from minus 450 to minus 850 ms.

The present paradigm clearly differs from iconic memory experiments in several ways. Most obvious is the markedly lower level of performance reported in Cumming and Coltheart (1969) and in the present studies. In addition the function of the probes in the present paradigm is to provide independent measures of item and position information. Clearly they do not operate as instructions to the scanner and processor producing the facilitatory partial report effect which is basic to the iconic memory paradigm.

While the task which confronts our subjects differs from that of the iconic memory experiment it is assumed that the storage processes at the subject's disposal do not differ. Sperling (1960) reports negligible change in performance out to unmasked exposure durations of 500 ms so it seems reasonable to conclude that no new storage facilities are available at the durations reported above. Both the partial independence of item and position information and the difficulty of formulating a plausible verbal mechanism to mediate position storage suggest that auditory-verbal STM is supplemented by a post-iconic visual code. If this visual code is to serve the acquisition and buffering function of Sperling's (1967) recognition buffer then its acquisition rate must be an order of magnitude greater than that of the verbal code. The superiority of position information at shorter exposures suggest that the visual code may lose item information while retaining position information (Cumming

and Coltheart's HIP). Item information to be preserved with stability must be acquired by the verbal code.

The distinction between the codes may be understood in terms of their indexing. Thus the visual code may be interrogated spatially or by identity whereas the verbal code may only be interrogated by name-identity.

References

ATKINSON, R. C. and SHIFFRIN, R. M. (1968). Human memory: a proposed system and its control processes. In SPENCE, K. W. and SPENCE, J. T. (Eds), *The Psychology of Learning and Motivation*. Vol. 2, pp. 89–195. New York: Academic Press.

AVERBACH, E. and SPERLING, G. (1961). Short-term storage of information in vision. In CHERRY, C. (Ed.), *Symposium on Information Theory*, pp. 196–211. London: Butterworth.

CUMMING, G. E. and COLTHEART, M. (1969). Short term memory: Are position and item information stored independently. *Psychonomic Science*, **15**, 216–8.

DICK, A. O. (1969). Relations between the sensory register and short term storage in tachistoscopic recognition. *Journal of Experimental Psychology*, **82**, 279–284.

HENDERSON, L., COLES, S., MANHEIM, M., MUIRHEAD, J. and PSUTKA, P. (1971). Orientation specific masking of letter features. *Nature*, **233**, 498–9.

JAMES, W. (1890). *The Principles of Psychology*. New York: Holt.

LISS, P. (1968). Does backward masking by visual noise stop stimulus processing? *Perception and Psychophysics*, **4**, 328–30.

NEISSER, U. (1967). *Cognitive Psychology*. New York: Appleton-Century-Crofts.

SCARBOROUGH, D. L. (1971). Memory for brief visual displays: the role of implicit speech. Paper presented to the Eastern Psychological Association, New York.

SPERLING, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, **74**, 1–29.

SPERLING, G. (1967). Successive approximations to a model for short term memory. *Acta Psychologica*, **27**, 285–92.

SPERLING, G. (1970). Short term memory, long-term memory, and scanning in the processing of visual information. In YOUNG, F. A. and LINDSLEY, D. B. (Eds), *Early Experience and Visual Information Processing in Perceptual and Reading Disorders*, pp. 198–218. Washington, D.C.: National Academy of Sciences.

THE MOVEMENT OF THE HAND TOWARDS A TARGET

W. D. A. BEGGS AND C. I. HOWARTH

Department of Psychology, University of Nottingham

Previous work on the approach trajectories of hands to targets in tracking and aiming tasks had produced contradictory evidence about the shape of these curves. This paper shows that these trajectories are a function of the level of practice of the subject; an interpretation in terms of the theory of intermittent control of movements is advanced. Previous contradictory data can be resolved by reference to the subjects' level of practice. In addition, this theory will encompass the differences between accurate and free movement trajectories reported earlier.

Introduction

A considerable body of data on the movement of a hand towards a target of some sort has been gathered. Woodworth (1899) distinguished an acceleration phase, a central phase of uniform velocity and a deceleration phase in movements of this type. Slower movements tended to have much longer deceleration phases, as did movements requiring accurate termination. Peters and Wenborne (1936) also found that the sizes of the acceleration and deceleration phases were dependent on the length of the movement, and the terminal precision required of the subject. Lowry, Maynard and Stegmerten (1948) claimed that approach curves showed a period of constant velocity. Taylor (1947) and Taylor and Birmingham (1948), using an oscilloscope and joystick technique, displayed position, velocity, acceleration and rate of change of acceleration on four separate oscilloscopes. They found:

(a) there was no period of constant velocity in approach movements;

(b) the relative sizes of the acceleration and deceleration phases depended on terminal accuracy: accurate movements tended to have longer deceleration phases, while movements of an approximate extent only had symmetrical patterns;

(c) when movement distance was increased, all the movement parameters increased in value, but quickly reached an asymptote as muscle forces reached their maximum;

(d) There were small variations in velocity and acceleration during the two main phases.

Vince (1948) confirmed the two phase nature of accurate movements in tracking tasks, with the deceleration phase of longer duration. Annett, Golby and Kay (1958) with their well-known high speed film technique found that most of the movement time was spent close to the target; they considered the approach as occurring in two distinct parts, a fast gross movement to the target area, followed by a slow terminal phase.

Murrell and Entwisle (1960) showed that approaches to targets were of a complex nature with changes in velocity during the main acceleration and deceleration phases. They suggested this was tremor: they found acceleration to occupy about one third of the movement time. Crossman and Goodeve (1963) also found irregularity in approach curves; they suggested that this was evidence for corrective responses. They also confirmed earlier work on accurate and "free" movements. Edwards (1965) showed that no period of constant velocity occurred, but confirmed that deceleration took longer than acceleration.

These workers used many sorts of task, including line drawing, aiming, step-tracking, joystick control, wrist rotation, and repetitive tapping. The results can thus be considered fairly general for approach movements. No particular use of this knowledge had been considered until Howarth, Beggs and Bowden (1971) measured the approach of the hand to a target with some accuracy. They used naive subjects, and found an almost symmetrical sigmoid relating time and distance to impact.

Beggs and Howarth (1972) also measured the approach curves of subjects who suffered interruption to their vision during aiming, and showed again almost symmetrical approach curves. From their theory of intermittent feed-back control of movement, they showed that terminal accuracy depended on the distance to impact, d_u, at which a final corrective movement could be applied to the movement, for both groups of subjects (Howarth *et al.*, 1971; Beggs and Howarth, 1972).

However, these naive subjects had apparently symmetrical approach curves, which is a contradictory result to the majority of previous experiments. The symmetry was both in the sizes of two main phases of movement and in the regularity across movement speed.

The main differences between these experiments and earlier work was the use of naive subjects and paced movements. In this paper, data on paced movement will again be reported. Initially, a naive subject was used, but given extended practice on an aiming task. This was monitored and the changes in approach to the target will be described.

Materials and Method

The apparatus used has been described in detail in Howarth *et al.* (1971). The movement made was similar to dart throwing; a pencil held in the normal fashion was moved from a home position near the subject's right shoulder to the vertically mounted target in front of the home position. The subject successively hit the base-plate and target coincident with metronome ticks: the movement distance was 508 mm.

A moveable infra-red beam which fell on a photocell was broken by the subject making repetitive movements to the target. This beam started a timer, which was stopped by the subject hitting the target. The infra-red beam and photo cell were placed at 10 equal intermediate distances between the baseplate and target.

Paper targets were used, and the subject was instructed to be accurate. Terminal accuracy was measured from the distribution of pencil marks around the vertical target line drawn on millimetre graph paper.

One subject was used for this experiment. She was an undergraduate psychology student, with corrected-to-normal vision. She took part on four successive days.

During each of the training sessions she performed 30 trials of 20 shots at the target. Three speeds were used, being 42, 85 and 125 ticks/min of a metronome or 1,425, 740 and 480 ms duration, at each of the 10 beam positions; these were presented randomly.

Results

These are shown in Tables I and II and in Figures 1, 2, 3 and 4. Successive sessions have been designated a, b, c and d.

TABLE I

Values of time to impact (ms) for different distances to impact, at three speeds and four levels of practice

Sessions Speed in t.p.m.	a			b			c			d		
	40	85	125	40	85	125	40	85	125	40	85	125
	0	0	0	0	0	0	0	0	0	0	0	0
50·8	135	88	68	371	119	72	382	170	77	549	192	79
101·2	255	153	100	467	196	108	507	260	135	662	326	160
152·4	410	288	135	462	264	146	569	311	157	664	356	185
Distance 203·2	515	256	172	544	323	182	668	367	182	713	370	231
(mm) 254·0	637	297	194	629	359	192	729	372	222	808	409	240
304·8	735	340	231	755	391	244	789	403	256	850	436	272
355·6	728	374	281	829	428	270	865	442	292	976	492	304
405·4	875	424	298	891	460	290	924	484	319	1015	497	316
457·2	972	483	333	951	499	331	992	519	352	1026	528	370
508·0	1425	740	480	1425	740	480	1425	740	480	1425	740	480

TABLE II

Values of mean square error in mm^2, and d_u in mm^2 at three speeds and four levels of practice

Speed in t.p.m.	a		b		c		d	
	E^2	d_u^2	E^2	d_u^2	E^2	d_u^2	E^2	d_u^2
40	2·27	11,653	2·00	1,008	1·35	645	1·08	413
85	11·85	57,006	8·60	30,272	3·73	16,129	3·35	7,458
125	19·12	158,014	13·43	153,006	16·83	137,522	7·61	115,845

On Figures 1, 2 and 3 is indicated the corrective reaction time of 290 ms as found by Beggs and Howarth (1970): in terms of their theory, corrections cannot be initiated after this time to impact.

Discussion

It is clear from inspection of Figures 1, 2 and 3 that the effect of practice is to change the shape of the approach curves to a target, both in the accelerative and particularly in the decelerative phase. We believe that the reason for this progressive change is that the subject was learning to minimize d_u, the distance through which the hand travels after the last corrective movement.

When naive, this subject had approach curves which were approximately symmetrical and did not vary in shape with movement time. This confirms our earlier work on naive subjects. However, as a result of practice she changed the shape of her approach curve so that she moved as close to the target as possible, as quickly as possible and then was able to spend more time on the terminal phase of the movement. This is obviously a more efficient strategy, and in terms of our
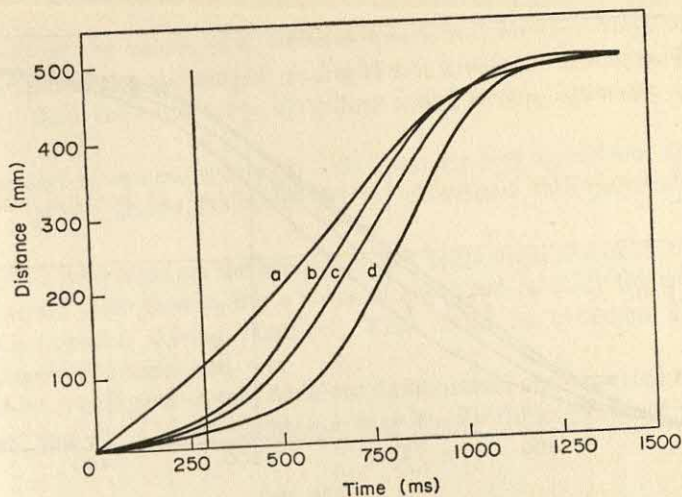
FIGURE 1. The smoothed relationship of distance to impact (d) and time to impact (t) for four levels of practice at 42 ticks per min.

theory, will reduce d_u to a smaller value than for the naive subject. At high speeds the approach curves of the practised subject are more symmetrical than at low speeds. This is presumably because of physical limitations on the acceleration and deceleration of the arm.

The apparent discrepancies in the literature on approach curves are easily resolved. All the studies which reported a change in the shape of the curve with speed used practised subjects. Our practised subject duplicated earlier results. Symmetrical curves are found for naive subjects or for free as opposed to aimed movement. For free movement it is not necessary to control the accuracy to any great extent so that there is no purpose in minimizing d_u. Symmetrical approach
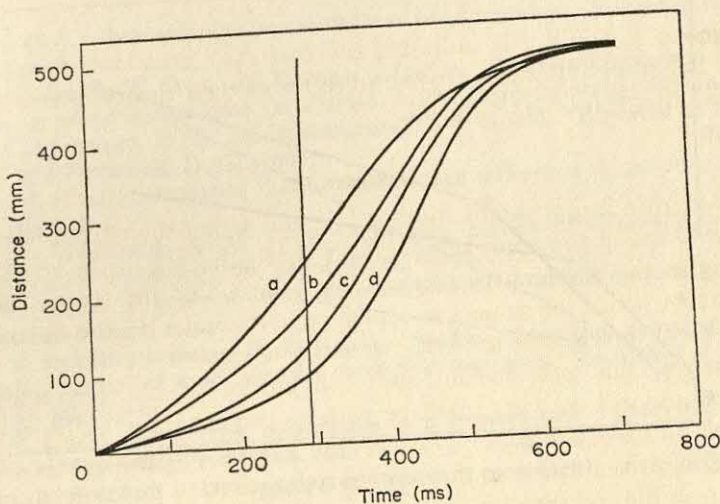
FIGURE 2. The smoothed relationship of distance to impact (d) and time to impact (t) for four levels of practice at 85 ticks per min.
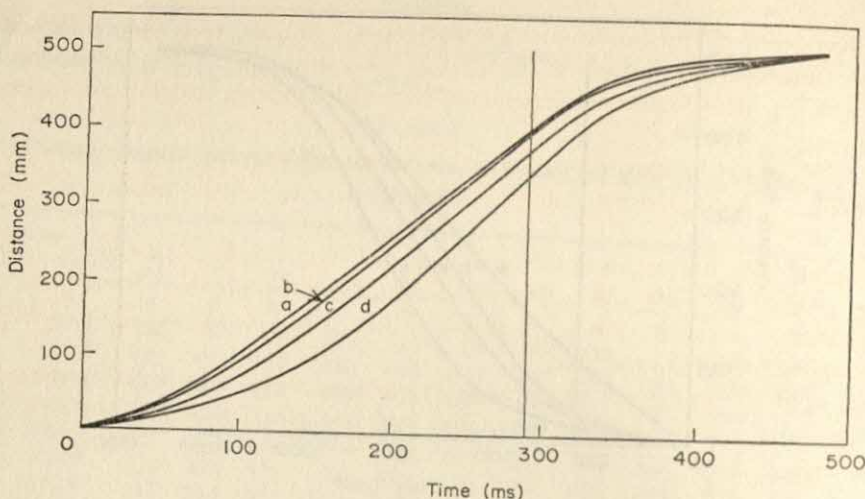
FIGURE 3. The smoothed relationship of distance to impact (d) and time to impact (t) for four levels of practice at 125 ticks per min.

curves are probably used because they minimize the muscle forces needed for a given extent of movement.

The effect of practice on accuracy provides an opportunity for a further test of our theory. The improvement in accuracy should be largely a result of the decrease in d_u, since error on target σ_ε should be predictable from the equation

$$\sigma_\varepsilon^2 = \sigma_0^2 + d_u^2 \sigma_\theta^2 \tag{1}$$

where σ_0^2 is the tremor variance

σ_θ^2 is the variance in the angular error of aiming.
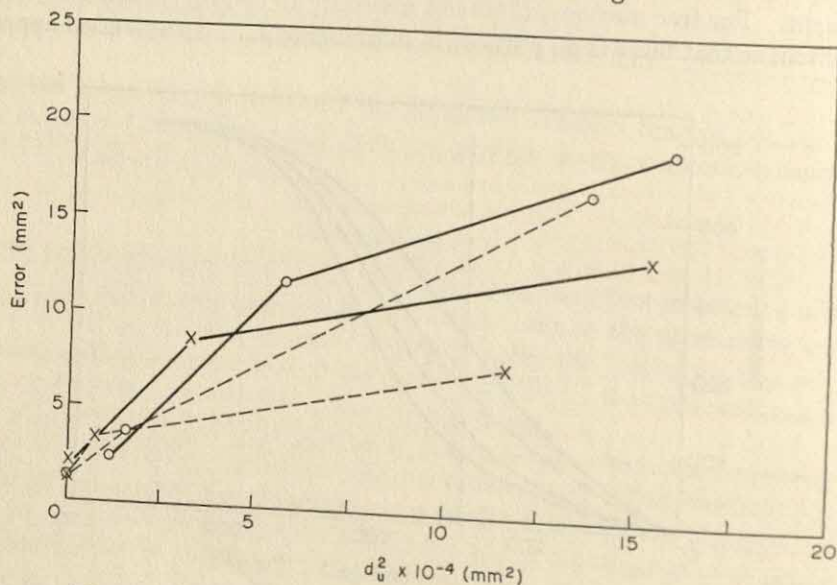
FIGURE 4. The relationship of error, in mm², to d_u^2 in mm² × 10^{-4}. ———O——— a; ———×——— b; ——— O ——— c; ——— × ——— d.

We have read off the values of d_u at each speed, and for each stage of practice, from the intercepts of the 290 ms line on Figures 1, 2 and 3. These values and the corresponding M.S. values of error have been fitted to this equation, and appear in Figure 4.

We have tended to assume that σ_θ is a constant for a given subject and a given movement. Figure 4 shows, however, that σ_θ decreased with practice, as well as d_u.

We believe the intercepts on the error axis, σ_0^2, to be measures of tremor. From Figure 4, it would seem that σ_0 has a value of about 1·6 mm for this subject, and remains fairly constant during practice. This could be expected if it were a physiological tremor component.

This is not a very elegant way in which to demonstrate our expected relationship between error and d_u; a better technique may clarify these effects of practice on σ_θ and σ_0^2.

References

ANNETT, J., GOLBY, C. W. and KAY, H. (1958). The measurement of elements in an assembly task—the information output of the human motor system. *Quarterly Journal of Experimental Psychology*, **10**, 1–11.

BEGGS, W. D. A. and HOWARTH, C. I. (1970). Movement control in a repetitive motor task. *Nature*, **225**, 752–3.

BEGGS, W. D. A. and HOWARTH, C. I. (1972). The accuracy of aiming at a target—some further evidence for intermittent control. *Acta Psychologica*. (In press.)

CROSSMAN, E. R. F. W. and GOODEVE, P. J. (1963). Feedback control of hand-movement and Fitts' Law. Unpublished communication to Experimental Psychology Society, Oxford.

EDWARDS, E. (1965). Research note on the study of hand motions. *Perceptual and Motor Skills*, **20**, 1098.

HOWARTH, C. I., BEGGS, W. D. A. and BOWDEN, J. (1971). The relationship between speed and accuracy of movement aimed at a target. *Acta Psychologica*, **35**, 207–18.

LOWRY, S. M., MAYNARD, H. B. and STEGMERTEN, G. J. (1948). Reported in MURRELL, K. F. H. and ENTWISLE, D. G. (1960).

MURRELL, K. F. H. and ENTWISLE, D. G. (1960). Age difference in movement pattern. *Nature*, **185**, 948.

PETERS, W. and WENBORNE, A. A. (1936). The time pattern of voluntary movements. *British Journal of Psychology*, **26**, 388–406; **27**, 60–73.

TAYLOR, F. V. (1947). A study of the acceleration pattern of manual corrective responses. *American Psychologist*, **2**, 340.

TAYLOR, F. V. and BIRMINGHAM, H. P. (1948). Studies of tracking behavior, II. The acceleration pattern of quick manual corrective responses. *Journal of Experimental Psychology*, **38**, 783–95.

VINCE, M. A. (1948). Corrective movements in a pursuit task. *Quarterly Journal of Experimental Psychology*, **1**, 85–106.

WOODWORTH, R. S. (1899). The accuracy of voluntary movement. *Psychological Review Monograph Supplement*, **3**, 13.

STUDIES ON THE FUNCTION OF
SIGHTING DOMINANCE

JEAN MONEY

Department of Psychology, The University of Western Ontario

The perceptual accuracy of the dominant and non-dominant eye was compared in four tachistoscopic experiments. The first two required that stimulus arrays be scanned during or immediately following stimulus presentations. The latter two tested digit recognition and dot localization in the absence of eye movement. Perceptual accuracy was significantly higher with the dominant eye than with the non-dominant in both scanning studies, while no perceptual differences were found between the eyes in the control experiments. In a fifth study, recordings of eye movement made during monocular scanning and searching tasks indicated that the dominant eye required significantly less time to complete the tasks than did the non-dominant. The results were interpreted as supporting Walls' (1951) theory that sighting dominance reflects asymmetrical motor functioning.

Introduction

The study of sighting dominance has largely been associated with the problem of laterality, and particularly that of crossed dominance (e.g. Harris, 1957). Little of the research has been directed towards the significance of sighting dominance *per se*, which is generally defined as a preference for one eye over the other in some form of visual alignment. Several early investigators (Miles, 1929, 1930; Schoen, 1936) outlined possible functions of the sighting eye, but there have been few attempts to organize their observations into a theory of sighting dominance.

An exception, however, has been a theory proposed by Walls (1951). Based primarily on the observation that the direction assigned to a point in space is the one stemming from the dominant eye, Walls has suggested that sighting dominance reflects an underlying asymmetry in motor functioning. Briefly, Walls outlined two ways in which sighting dominance could be associated with motor functioning. His first suggestion was that only the dominant eye participates in the hypothesized comparison and cancellation of efferent outflow with the subsequent reafferent stimulation during voluntary movement (Ludvigh, 1952; Von Holst, 1954). According to this view, the dominant eye would be primarily responsible for stability of the visual field during eye movement and also for the control or guidance of ocular movement. Walls' second hypothesis was that the dominant eye may initiate the muscular adjustments involved in fixation, with the non-dominant eye making corresponding reflexive adjustments to maintain binocular fusion.

If sighting dominance is associated with an asymmetry in motor functioning, then some consequences of this asymmetry should be detected in monocular vision. This paper reports a comparison of the perceptual accuracy of the dominant and

non-dominant eyes on four tasks which varied in their degree of probable motor involvement.   According to Walls' theory, the non-dominant eye may be expected to show a deficit, relative to the dominant, in situations in which accuracy of perception would be enhanced by rapid eye movement, but not in the absence of movement.   Therefore, the first two experiments required that stimulus arrays be rapidly scanned, while in the latter two, eye movement was not required.   In addition, monocular eye movement was photographically recorded during a series of scanning operations.

General Method

Subjects

Subjects in all experiments were paid volunteers from the University of Western Ontario. Age range was 16–29 years.   To obtain the desired number of subjects in each sighting classification and an approximately equal distribution of males and females, information on sighting dominance was collected on a large number of subjects while they were participating in a variety of other experiments.   Then in each of the following studies, some subjects were selected at random and the rest on the basis of their known sex and sighting dominance.

Sighting and acuity measures

Sighting dominance was determined from Miles' (1929, 1930) A-B-C test and two pointings.   In the latter, the subject aligned first the tip of his preferred index finger and then the end of a blackboard pointer held in both hands against his nose, with a wall target approximately 15 ft (1 ft = 0·3048 m) in front of him.   Subjects who did not exhibit a clear sighting preference were classified as ambi-eyed and their data excluded from the analysis.   Furthermore, only the data from subjects with at least 20/30 (corrected) vision in each eye, as measured on the Snellen chart, were included in the analysis.   Of the 89 subjects tested for these experiments, 11 were classified as ambi-eyed and one was rejected for insufficient acuity.

Apparatus

In the first four experiments, stimuli were presented by means of a Gerbrands Harvard two-field mirror tachistoscope (Model T-2B-1).   With this type of tachistoscope, the subjects' eyes are approximately 2 ft from the exposure field.   Illumination was set at the maximum level of approximately 8 ft lamberts (1 ft lambert = 3·42626 cd/m²) in both the pre-exposure and exposure fields.   The tachistoscope was adapted for monocular viewing by insertion of a black cardboard into the eye piece in front of one eye.

In Experiment V, horizontal eye movements were recorded by means of Biometrics Reading Eye II, which will be described in greater detail in a later section.

Data analysis

In the tachistoscopic experiments, the score for each eye was the number of correct responses (digits identified or dot localizations).   The eye movement recordings were scored for the total time required by each eye to complete the tasks.   A three-way analysis of variance incorporated sex (male and female) and sighting dominance (left and right) as between-subject variables and viewing condition (dominant, non-dominant and binocular) as a within-subject variable.   (For purposes of comparison, a binocular condition was included in the tachistoscopic studies.)   The main effects of sex and sighting dominance did not reach significance in any study and will not be referred to again.

Experiment I

Walls' theory would predict more efficient functioning with the dominant eye than with the non-dominant in situations requiring rapid initiation and control of

ocular movements. Therefore, if subjects were required to scan monocularly, stimulus arrays presented for a limited time, one would expect more of the stimulus material to be correctly reported with the dominant eye.

Subjects

Nine left-eye dominant (4 male and 5 female) and 18 right-eye dominant subjects (10 male and 8 female) participated.

Procedure

Each stimulus array, consisting of three digits, was presented tachistoscopically for 500 ms. One digit appeared at the midpoint of each of three of the four sides of the 7·7 in (195 mm) square exposure field, ⅜ in (10 mm) from its outer edge (see Fig. 1). The four positions in
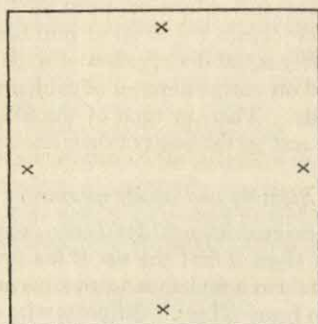
FIGURE 1.

the field were used equally often; otherwise, the combination of positions in each stimulus array was randomly determined. The visual angle subtended by the two digits on opposite sides was approximately 16°. There was no pre-exposure fixation point.

Twelve different arrays were made from Letraset digits 2–9. Each digit measured $\frac{1}{16}$ in. (1·5 mm) in height, subtending a visual angle of approximately 9' of arc. No digit appeared more than once in an array; otherwise, the selection of digits was randomly determined.

Size of the digits precluded recognition in peripheral vision. The requirement that each digit receive foveal placement ensured movement of the eyes during the stimulus presentation.

Identical copies of the 12 stimulus arrays were presented to each eye monocularly and binocularly in a single test session. Trials were presented in a fixed random block design, each block consisting of one binocular trial and one monocular trial with each eye. Presentation of arrays within each viewing condition was in a fixed random order. Each trial was preceded by a ready signal. Subjects reported the digits verbally following each presentation.

Results

The maximum score for each viewing condition was 36. Analysis of variance indicated a significant difference among the three conditions ($F = 21·7$, $df = 2, 46$, $P < 0·001$). Subsequent t-tests indicated that the number of digits correctly reported with the dominant eye (mean of 16·3) was significantly greater than with the non-dominant (mean of 14·8; $t = 3·2$, $df = 46$, $P < 0·01$). Also, scores in binocular viewing (mean of 18·0) were significantly higher than with either eye alone ($t = 3·8$, $df = 46$, $P < 0·001$ for the sighting eye and $t = 7·1$, $df = 46$, $P < 0·001$ for the non sighting eye).

Experiment II

Experiment I demonstrated that when stimulus arrays were briefly presented for scanning, perceptual accuracy was significantly higher when viewing with the dominant eye than with the non-dominant. However, it is possible that the effect of an asymmetry in motor functioning may not be limited to eye movement made during the stimulus presentation. Previous studies have demonstrated that perception can be influenced by post-exposure processes. Crovitz (1962), for example, found that perceptual accuracy was higher in the visual field which corresponded to the direction of an eye movement made during the post-exposure period.

In this experiment, subjects were required to make a post-exposure scanning movement toward the stimuli. If a difference does exist between the eyes in their efficiency of motor functioning, then this may result in their stimulus traces receiving an asymmetrical amount of post-exposure facilitation. Heron (1957) has suggested that such facilitation comes primarily from eye movement centres. Therefore, one could expect a perceptual deficit with the non-dominant eye even when the motor involvement was restricted to the period following stimulus presentation.

Subjects

Eight left-eye dominant (4 male and 4 female) and 19 right-eye dominant subjects (9 male and 10 female) participated.

Procedure

The stimulus arrays, consisting of three digits were presented for 100 ms, precluding voluntary eye movement during the stimulus presentation. Dodge, in 1908, reported the mean latency for a saccade to be approximately 200 ms (in Woodworth and Schlosberg, 1954). The pre-exposure field contained a dot in the centre to indicate the fixation point. The three digits were located 1·2 in. (30 mm), 2·4 in. (60 mm) and 3·6 in. (90 mm) from fixation in a straight line which extended vertically above or below the dot or horizontally to its left or right (see Fig. 2). The post-exposure movement, therefore, was made in one
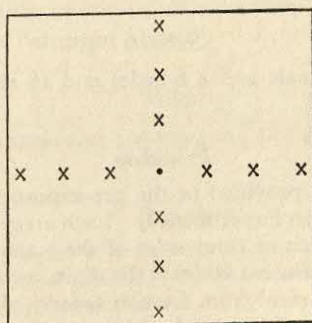
FIGURE 2.

of four directions from the central dot toward the stimuli. The visual angle subtended between the fixation point and the outer digit was approximately 8°.

Twelve different stimulus arrays, employing each direction equally often, were made from Letraset digits 2–9. Each digit measured $\frac{3}{16}$ in. (4·5 mm) in height, subtending a visual angle of approximately 27′ of arc. The combination of digits appearing in each array was

the same as in Experiment I. Identical copies of the 12 stimulus arrays were presented to each eye monocularly and binocularly in the same way as described for Experiment I.

Subjects were instructed to fixate the central dot until the stimuli appeared and then to scan quickly the row of digits. The importance of looking toward the location of the outer digit, even if it was no longer visible, was emphasized. It was thought that positioning the digits in a straight line might maximize post-exposure facilitation since all the digits would be scanned by one movement. Initial fixation was ensured by giving a ready signal before each presentation. Subjects verbally reported the digits following each trial.

Results

The maximum score for each viewing condition was 36. Analysis of variance indicated a significant difference among the three conditions ($F = 8.8$, $df = 2, 46$, $P < 0.001$). Subsequent t-tests indicated that significantly more digits were correctly reported with the dominant eye (mean of 19·3) than with the non-dominant (mean of 18·1; $t = 2.8$, $df = 46$, $P < 0.01$). The mean score in binocular viewing (19·7) did not differ significantly from that for the dominant eye ($t = 1.1$, $df = 46$, $P > 0.10$) but was significantly higher than that for the non-dominant eye ($t = 3.9$, $df = 46$, $P < 0.001$).

Experiment III

The perceptual superiority of the dominant eye during scanning and post-exposure scanning supports the hypothesis of an asymmetry in motor functioning. In terms of Walls' theory, there should be no perceptual difference between the eyes if motor functioning is kept to a minimum. The next two experiments, therefore, were designed to compare the perceptual performance of the two eyes when the probability of motor involvement was minimized by employing stimulus exposure durations which precluded eye movement during the presentation, and by instructing against post-exposure activity.

Experiment III tested digit recognition under these conditions.

Subjects

Eight left-eye dominant (4 male and 4 female) and 18 right-eye dominant subjects (9 male and 9 female) participated.

Procedure

A central fixation point was provided in the pre-exposure field. The stimulus arrays were similar to those described in Experiment I. Each array consisted of three digits, with one digit at the midpoint of each of three sides of the square exposure field. One major difference from the former experiment was that the digits were moved in from the periphery of the field to only 1 in. (25·4 mm) from fixation (nearly 2½° of visual angle; see Fig. 3). The proximity of the digits to fixation plus the use of larger Letraset digits measuring $\frac{2}{16}$ in. (3 mm) in height, eliminated the necessity of scanning the arrays. The combination of positions and digits in each array was the same as in Experiment I.

A second important difference from the first experiment was that the length of stimulus presentation was reduced to 100 ms. Also, subjects were instructed to fixate on the central dot during the post-exposure period. If a tendency to look toward the digits existed during this period, the positioning of the digits in three different directions from fixation should have minimized any facilitation which might have resulted from a post-exposure tendency.
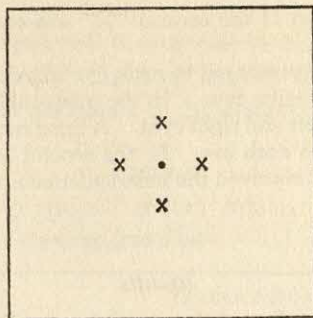
FIGURE 3.

Trials were presented in the same order as Experiment I. Each presentation was preceded by a ready signal to ensure central fixation. Subjects verbally reported the digits following each trial.

Results

Maximum possible score for each viewing condition was 36. Analysis of variance indicated a significant difference among the three conditions ($F = 16\cdot3$, $df = 2, 44$, $P < 0\cdot001$). However, subsequent t-tests indicated that the number of digits correctly reported with the dominant eye (mean of $18\cdot6$) did not differ significantly from that with the non-dominant (mean of $18\cdot4$; $t = 0\cdot1$, $df = 44$, $P > 0\cdot10$). The variability in this factor resulted from significantly higher scores in binocular viewing (mean of $22\cdot6$) than for either the dominant eye ($t = 4\cdot8$, $df = 44$, $P < 0\cdot001$) or the non-dominant alone ($t = 5\cdot0$, $df = 44$, $P < 0\cdot001$).

Experiment IV

The fourth study compared the two eyes in the spatial localization of dots. Voluntary eye movement during the stimulus presentation was precluded by a 10-ms exposure duration. In the absence of eye movement toward the stimuli, the basis for judging the location of the dot would simply be the location of its retinal image relative to the eye's "straight ahead".

Subjects

Eight left-eye dominant (4 male and 4 female) and 18 right-eye dominant subjects (9 male and 9 female) participated.

Procedure

A shortened version of Kimura's (1969) dot location task was presented. The pre-exposure field consisted of a circle 5 in. (127 mm) in diameter, with an "X" in the centre to indicate the fixation point. Each stimulus was a single dot presented in the circle for 10 ms. The subject reported the position of the dot from a matrix card containing all the positions presented.

Letraset dots were presented at 1 in. (25·4 mm), 1·5 in. (38 mm) or 2 in. (51 mm) from fixation in 24 of the 48 positions used by Kimura. Four dots at each distance were selected at random from the left visual field and their mirror images from the right field. The response matrix, containing the original 48 positions plus two practice positions, was mounted on the tachistoscope just over the subject's head, so that he had to look out of the apparatus

to make his response. Fixation of the central "X" was ensured by giving a ready signal before each presentation.

The same 24 positions were presented to each eye monocularly and binocularly. Each subject was tested on two successive days. In the monocular session, trials were presented in a fixed random order to the left and right eyes. A fixed random order was also used in the presentation of the positions to each eye. In the second session, viewing was binocular. Half of the subjects of each sex received the monocular condition first, the rest received the binocular first.

Results

Maximum possible score for each viewing condition was 24. Analysis of variance indicated a significant difference among the three conditions ($F = 15\cdot8$, $df = 2, 44, P < 0\cdot001$). However, a subsequent t-test indicated that the number of dots correctly localized with the dominant eye (mean of $10\cdot53$) did not differ significantly from that with the non-dominant (mean of $10\cdot57$; $t = 0\cdot07$, $df = 44$, $P > 0\cdot10$). The variability in this factor resulted from significantly higher scores in binocular viewing (mean of $13\cdot1$) than with either the dominant eye ($t = 4\cdot9$, $df = 44, P < 0\cdot001$) or the non-dominant eye alone ($t = 4\cdot8$, $df = 44, P < 0\cdot001$). The results of the four tachistoscopic studies are summarized in Table I.

TABLE I
Mean number of correct responses for each viewing condition in Experiments I–IV

	Experiment			
	I	II	III	IV
Viewing condition	Scanning	Post-exposure scanning	Digit recognition	Dot localization
	(maximum = 36)	(maximum = 36)	(maximum = 36)	(maximum = 24)
Non-dominant eye	14·8	18·1	18·4	10·57
Dominant eye	16·3	19·3	18·6	10·53
Binocular	18·0	19·7	22·6	13·1

Peripheral versus central acuity

In the scanning experiment, the stimuli were located in the periphery of the visual field and a difference in perceptual accuracy was found between the dominant and non-dominant eye. In digit recognition, the stimuli appeared close to the fixation point and no difference was found. To determine if the more accurate perception with the dominant eye resulted from better peripheral vision, rather than from facility in eye movement, the two experiments in which the stimuli varied in distance from fixation were reanalyzed in terms of the number of correct responses at each distance from fixation in monocular viewing. If the two eyes differ in peripheral acuity, one would expect an interaction between distance from fixation and the viewing eye.

The results of the reanalysis for post-exposure scanning and dot localization are shown in Tables II and III. In both studies accuracy was highest for the stimuli closest to fixation and decreased as the distance from fixation increased.

TABLE II

Mean number of digits correctly reported at each distance from fixation in post-exposure scanning (maximum = 12)

	Distance from fixation		
	Near (2°, 52′)	Middle (5°, 43′)	Far (8°, 32′)
Viewing eye			
Dominant eye	9·6	6·7	2·9
Non-dominant eye	9·7	5·9	2·6

TABLE III

Mean number of dots correctly localized at each distance from fixation (maximum = 8)

	Distance from fixation		
	Near (2°, 23′)	Middle (3°, 35′)	Far (4°, 47′)
Viewing eye			
Dominant eye	4·3	3·3	2·9
Non-dominant eye	4·0	3·8	2·8

In post-exposure scanning, analysis of variance indicated significant main effects for both Viewing Condition ($F = 8·4$, $df = 1, 23$, $P < 0·01$) and Distance from fixation ($F = 180·7$, $df = 2, 46$, $P < 0·001$). However, the interaction between Viewing Condition and Distance failed to reach a significance ($F = 3·1$, $df = 2$, 46, $P > 0·10$). In dot localization, again there was significant main effect for Distance from fixation ($F = 6·1$, $df = 2, 44$, $P < 0·01$), but the interaction between Viewing Condition and Distance did not reach significance ($F = 1·4$, $df = 2, 44$, $P > 0·10$).

For both eyes, the decline in accuracy as distance from fixation increased was constant. The data do not suggest that the dominant eye has higher acuity than the non-dominant specifically in *peripheral* vision, a finding consistent with reports that sighting dominance is not associated with acuity in central vision either (Gahagan, 1933).

Experiment V

Although this research was concerned primarily with the perceptual consequences of the hypothesized motor asymmetry, some preliminary data obtained from recording monocular eye movement provides additional support for Walls' theory. The results are congruent with the perceptual deficit found for the non-dominant eye in the scanning experiments, but it must be emphasized that this study was intended only as a preliminary attempt to compare monocular eye movement and more detailed study is required.

If the efficiency of motor functioning does differ in some way between the eyes, then one might expect a difference in the overall pattern of speed of monocular eye movements. This experiment was designed to measure the speed of the two eyes during visual scanning, tracing and searching.

Subjects

Eight left-eye dominant (4 male and 4 female) and 10 right-eye dominant subjects (5 male and 5 female) participated.

Procedure

Horizontal eye movements were recorded by means of Biometrics Reading Eye II. This apparatus monitors the light reflected from the corneas while the eyes are directed towards a small viewing stand which holds the 3½ in. (89 mm) × 5 in. (127 mm) stimulus cards. The record of movement made by each eye is printed onto heat-sensitive paper, so that an immediate record of the movement is visible.

Monocular movement was recorded during eight discrete trials. These included three types of scanning tasks (see Fig. 4). Four trials required rapid changes in fixation between
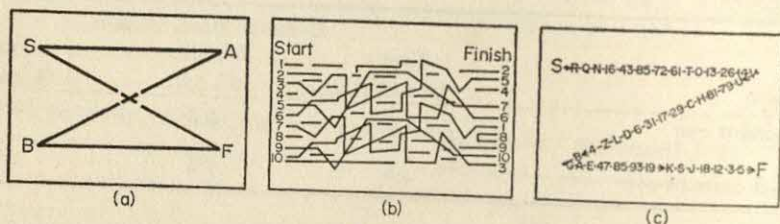
FIGURE 4.

designated letters on the stimulus card [Fig. 4(a)]. For example, subjects scanned from "A" to "S", or from "X" to "S" to "F", etc. Two trials involved tracing a line from a digit on the left side of the card, through a maze of lines, to a different digit on the right [Fig. 4(b)], and two trials required searching for a designated symbol along a string of digits and letters arranged in the form of a "Z" [Fig. 4(c)].

Preceding each trial, the required movements were explained to the subject who then fixated on the appropriate symbol until the experimenter said "go". The recording device was activated simultaneously with the verbal command. Subjects were instructed to make the sequence of movements as rapidly as possible and to indicate completion of each task by eyelid closure immediately upon sighting the required symbol. The time elapsed was measured on each trial from the beginning of the record until eyelid closure, by counting the number of millimeters (100 ms units) crossed by the recording pen.

An ocular, rather than a verbal or manual response, was selected to indicate the termination of each movement sequence because such a response would appear directly in the eye movement record and, therefore, be the most accurate and sensitive response indicator.

Both the left and right eyes were tested on the same tasks following a binocular practice session in which similar tasks were presented. Order of presentation ensured that practice effects were counterbalanced across eyes. Commercial eye patches were worn during the monocular recordings.

Results

The score for each eye was the total time summed over its eight trials. Analysis of variance indicated that the non-dominant eye (mean of 18·7 s) took significantly

longer to complete the tasks than did the dominant (mean of 17.5 s; $F = 7.8$, $df = 1, 14$, $P < 0.02$). Although the ease of voluntary eyelid closure may be asymmetrically facilitated by winking or sighting, this potential bias cannot be invoked as an explanation of these results. Subjects were instructed to work with maximum speed and it was assumed that they closed both eyelids together, since no facial contortions were observed to indicate otherwise. Crider's (1941) data would in any case suggest that the non-dominant eye is easier to close than the dominant and, indeed, the dominant eye has sometimes been defined as the eye which is harder to close alone (Walls, 1951). Therefore, if there was differential speed of eyelid closure it should have contributed to shorter times for the non-dominant eye, rather than the dominant.

Discussion

The results of these experiments support Walls' (1951) theory that sighting dominance is associated with asymmetrical motor functioning. In situations requiring rapid motor involvement, a perceptual deficit was demonstrated by the non-dominant eye. Thus, perceptual accuracy was found to be significantly higher when viewing with the dominant eye, both when stimulus arrays were briefly presented and scanned, and when scanning was restricted to the post-exposure period. However, there was no evidence of a perceptual deficit with the non-dominant eye, either in digit recognition or spatial localization, when scanning was not required.

In general, binocular viewing was found to be superior to either eye alone. The magnitude of the difference between binocular vision and the dominant eye was larger in the control studies than in those which involved scanning (Fig. 1), suggesting that vision is primarily a function of the dominant eye when rapid changes in fixation are required. In fact, the difference between binocular vision and the dominant eye was not significant in post-exposure scanning. In the other experiments, the larger visual field and higher acuity associated with binocular vision could account for its perceptual superiority, particularly in the control studies in which a single fixation was maintained.

The results do not permit an evaluation of Walls' specific hypotheses concerning the nature of the motor asymmetry. Presumably, a difference in the speed of movement initiation between the eyes could be detected in a binocular record of eye movement. Such attempts with the Reading Eye II, however, were unsuccessful, possibly because this apparatus is insensitive to eye movement during intervals of less than 100 ms. The monocular records demonstrated a post-initiation difference in motor ability, and supported the general trend in this research that the major manifestation of the asymmetry is inequality between the eyes in their speed of motor functioning.

References

CRIDER, B. (1941).   Eye-closure facility and eye dominance.   *Journal of Genetic Psychology*, **58**, 425–6.

CROVITZ, H. F. (1962).   Tendencies to eye movement and perceptual accuracy.   *Journal of Experimental Psychology*, **63**, 495–8.

GAHAGAN, L. (1933).   Visual dominance-acuity relationships.   *Journal of General Psychology*, **9**, 455–9.

HARRIS, A. J. (1957).   Lateral dominance, directional confusion and reading disability.   *Journal of Psychology*, **44**, 283–94.

HERON, W. (1957).   Perception as a function of retinal locus and attention.   *American Journal of Psychology*, **70**, 38–48.

HOLST, E. VON (1954).   Relations between the central nervous system and the peripheral organs.   *British Journal of Animal Behaviour*, **2**, 89–94.

KIMURA, D. (1969).   Spatial localization in left and right visual fields.   *Canadian Journal of Psychology*, **23**, 445–58.

LUDVIGH, E. (1952).   Control of ocular movements and visual interpretation of environment.   *AMA Archives of Ophthalmology*, **48**, 442–8.

MILES, W. R. (1929).   Ocular dominance demonstrated by unconscious sighting.   *Journal of Experimental Psychology*, **12**, 113–36.

MILES, W. R. (1930).   Ocular dominance in human adults.   *Journal of General Psychology*, **3**, 412–30.

SCHOEN, Z. J. (1936).   Functional asymmetry within the visual apparatus.   *American Journal of Optometry*, **13**, 130–4.

WALLS, G. L. (1951).   A theory of ocular dominance.   *AMA Archives of Ophthalmology*, **45**, 387–412.

WOODWORTH, R. S. and SCHLOSBERG, H. (1954).   *Experimental Psychology*.   New York: Holt, Rinehart & Winston.

SPECIFICITY AND SAVINGS OF BEHAVIOURAL HABITUATION OVER A SERIES OF INTRA- AND INTER-MODAL STIMULI

SANDRA E. FILE† AND I. STEELE RUSSELL

MRC Unit for Neural Mechanisms of Behaviour, Department of Psychology, University College, London

A series of four tone stimuli, differing either in tonal frequency or in melody, was presented at the rate of one presentation per day to a group of rats. Habituation was specific to the frequency of simple auditory stimuli and to the melody of patterned stimuli, as shown both by the distraction to stimulus change, and by the trials to habituate to the new stimulus. Savings in habituation occurred when a series of tone stimuli was presented, but was not found for a series of inter-modal stimuli.

Introduction

Although early experimenters on behavioural habituation regarded it as a simple form of learning (Harris, 1943) they did little to support this assumption empirically. Many of the physiological investigations of habituation highlighted the difficulties in distinguishing habituation from fatigue, e.g. the large number of stimulus presentations and the rapid repetition rates needed to obtain response decrement (e.g. Wickelgren, 1967; Sharpless and Jasper, 1956). Such findings led to a serious questioning of whether habituation was in fact a form of learning (Horn, 1967).

Recent developments have suggested a return to the notion that habituation should be regarded as a type of learning (Miller, 1967; Kimmel and Goldstein, 1967; Bishop and Kimmel, 1969). Some empirical support for this view comes from the studies of Barrett and Ray (1969), who concluded that the acquisition of habituation involves processes similar to the acquisition of a learned response. Blanchard, Shelton and Blanchard (1970) obtained evidence that habituation involves information storage, and it is therefore unlikely that it depends on any short-term change, such as fatigue. The critical parameters in assessing whether habituation is a form of learning are not well defined, but must involve principally the specificity and duration of the behavioural changes following stimulus repetition.

Habituation involves the registration and storage of some of the stimulus features of the environment, since an animal habituated to one situation is immediately

†Present address: Psychology Department, City of London Polytechnic.

distracted if the situation is changed. There have been several reports that habituation is specific to stimulus frequency, but these have been based on physiological measures (Sharpless and Jasper, 1956; Sokolov 1960; Zimny and Schwabb, 1965; Fokkema, 1966; Corman, 1967; Meyers and Gullikson, 1968). The physiological measures used are purported to reflect the mechanisms involved in behavioural habituation, i.e. cessation of orienting to a stimulus, but the acceptability of this presumption can only rest on a validation of these findings at a behavioural level. Only the latter can provide a direct assessment of the specificity of the behavioural changes which occur in habituation.

One report using a behavioural measure of habituation has suggested that it is specific to stimulus frequency (Miller and Murray, 1966). Animals were habituated to one tonal frequency and then the orienting responses measured to a variety of frequency values. Any change in frequency produced a return of orienting and there was some evidence that tones more distant from the habituated one produced a greater response. The order of presentation of test frequencies was not counterbalanced and it is possible that these results were due to the order of testing, rather than to the specificity of habituation.

The present experiment involved an improved version of the Miller and Murray technique. The length of disruption of a rat's on-going behaviour, produced by the presentation of a stimulus, was taken as a measure of the orienting response. Drinking was selected as it was an easily established on-going behaviour and had a steady rate of seven laps per second (Stellar and Hill, 1952).

The aims of the first experiment were to investigate at a behavioural level: (1) the degree of tonal specificity of habituation to simple auditory stimuli; (2) the degree of melodic specificity of habituation to patterned auditory stimuli; (3) the amount of savings in habituation over a series of different stimuli within the auditory modality.

Experiment I

Method

Subjects

Forty-eight male hooded rats, 300–350 g in weight, were housed individually in a 12 h light–12 h dark cycle. They were subjected to 72 h water deprivation and thereafter had daily access to water in the test chamber and immediately following an experimental session, in sufficient quantity to maintain a constant body weight. At all times food was available ad libitum in the home cages.

Apparatus

The test chamber was 20·5 cm in height with a grid floor 34 × 18 cm, enclosed in an acoustically insulated refrigerator shell. The rat had access to a water spout through a slit, 3·5 × 1 cm, in the middle of the end wall of the chamber. A drinkometer was used to record the rat's licking and all the experimental events were automatically programmed, using standard relay equipment. The source of tones was a Weinbridge oscillator which fed into a Sinclair amplifier and then to a loudspeaker (Wharfedale Super 3 Tweeter type), positioned on the lid of the test chamber.

Procedure

Pre-habituation drinking sessions were given in the test chamber for five days, by which time the rat would start licking immediately on entry to the chamber, and would continue without significant pause for at least 5 min. From the sixth day onwards the two-hundredth lick switched on a control period of 20 s in which the number of licks occurring was recorded. Following this period the next two-hundredth lick switched on a tone stimulus of 77·5 dB for 20 s, and the number of licks occurring during this time was measured. Only one tone presentation per day was given.

The rats were randomly assigned to eight groups each of which received a series of four tone stimuli. The order of stimulus presentation differed for each group and is given in Table I.

TABLE I

Order of stimulus presentation to each group

| Experimental group | Stimulus positions | | | |
	1	2	3	4
A1	7	9	C	R
A2	7	9	R	C
B1	9	7	C	R
B2	9	7	R	C
C1	C	R	7	9
C2	C	R	9	7
D1	R	C	7	9
D2	R	C	9	7

The tones used were 7 kHz, 9 kHz, a complex tone (C), and a reversed complex tone (R). C consisted of a pattern of 9 kHz for 5 s, succeeded by 7 kHz for 10 s and followed by 9 kHz for 5 s. R was a pattern of 7 kHz for 5 s, 9 kHz for 10 s and then 7 kHz for 5 s. These frequencies were chosen since they are of the same subjective intensity to the rat (Clack and Harris, 1963; Gourevitch and Hack, 1966), and hence a change in frequency would not also involve a change in apparent intensity.

In all cases the same stimulus was presented until the criterion (no interruption in the rat's licking) was reached on two successive days. The following day the next stimulus was presented. This continued until the animal had habituated to all four of the stimuli in the series.

Results

Interpretations about the frequency specificity of habituation were made on the basis of the distraction caused by the presentation of a stimulus of a different frequency from the habituated one. It was therefore necessary to demonstrate that none of the stimuli was intrinsically more distracting than the others, in spite of being equal in subjective intensity. The measure of distraction was the difference between the number of licks in the control and tone periods, divided by the number of licks in the control period. This distraction ratio was calculated for the first stimulus presentation to each group on the first day of training. In addition, the number of trials needed to habituate to each stimulus was recorded. The

criterion of habituation was a ratio of ≤0·05. These results are shown in Table II. Analysis of variance showed that there was no significant difference between any of these scores ($F = 0·38$, $df = 1$, 22, $P > 0·10$ for the greatest difference, i.e. trials to criterion for C and R tones). Therefore the assumption can be made that all the stimuli were equally distracting, as reflected by both the length of the initial pause and the number of trials required to habituate.

In a previous experiment (File, 1969) it was found that once two successive criterion trials were made, the animal remained habituated on the third and fourth trials. In this experiment the control data, i.e. the distraction ratio to an un-changed stimulus, was therefore obtained from the second criterion trial.

TABLE II

Distraction resulting from the initial stimulus presentation

	Stimulus			
	7 kHz	9 kHz	C	R
Distraction ratio	0·64	0·72	0·68	0·68
Trials to criterion	6·9	6·0	7·75	6·0

(a) *The degree of tonal specificity to simple auditory stimuli*

Table III gives the distraction resulting from the first change in stimulus frequency. For groups A1 and A2 the change was from the habituated frequency of 7 kHz to the novel one of 9 kHz. This produced a return of orienting with a mean distraction ratio of 0·11, and 9/12 of the animals in this group had distraction ratios greater than the criterion level. This ratio was significantly greater than the control distraction to an unchanged stimulus ($t = 3·01$, $df = 11$, $P < 0·01$, one-tailed test). For groups B1 and B2 the change was from the habituated fre-quency of 9 kHz to the novel one of 7 kHz. This change produced a mean dis-traction ratio of 0·21, with 8/12 of the animals in this group having ratios greater than the criterion level. This was also significantly greater than the control level ($t = 2·93$, $df = 11$, $P < 0·01$, one-tailed test). It took the groups a mean of 3·8 and 2·8 trials respectively to habituate to this new stimulus of 9 or 7 kHz. These results demonstrate tonal specificity of habituation to simple auditory stimuli.

TABLE III

Distraction resulting from the first stimulus change

	Stimulus change			
	7–9 kHz	9–7 kHz	C–R	R–C
Distraction ratio	0·11	0·21	0·24	0·18
Trials to criterion	3·8	2·8	3·5	2·3

(b) *The degree of melodic specificity to patterned auditory stimuli*

Table III also shows the distraction resulting from a change in the melody of the habituated stimulus i.e. a change from C to R, or vice versa. This change in

melody resulted in distraction ratios of 0·24 and 0·18 for the two groups respectively, and in both cases 9/12 animals had distraction ratios greater than the criterion level. These ratios were significantly greater than the control levels ($t = 3·08$, $df = 11$, $P < 0·01$ and $t = 2·51$, $df = 11$, $P < 0·025$ respectively on one-tailed tests). It took the groups a mean of 2·3 and 3·5 trials to habituate to the new stimulus, which only represented a change in melody.

(c) *Savings in habituation over a series of tone stimuli*

Savings in habituation over a series of stimuli would be reflected in a decrease in the number of trials needed to habituate to each new stimulus, even though each might still produce some return of orienting. Figure 1 shows the trials needed to habituate to each of the four stimuli, over stimulus presentation orders 1–4. Each point on the graph represents the mean score of two groups (i.e. 12 animals). In each case the two groups combined received the same stimulus, preceded by the same number of prior stimuli (see Table I).
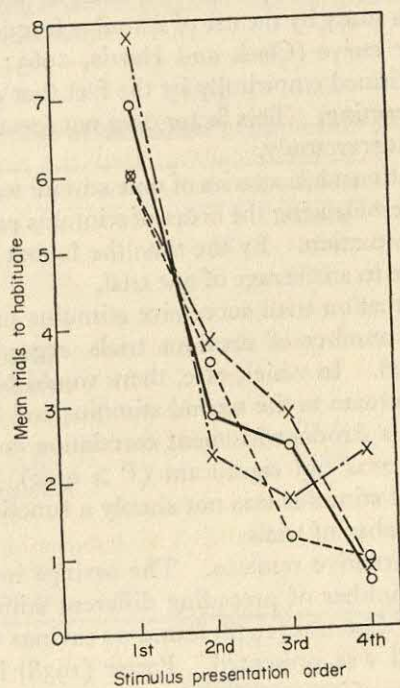
FIGURE 1. Mean trials to reach criterion for stimulus presentations 1–4. O————O 7 kHz; X— — — — —X 9 kHz; X————————X C; O———————O R.

In all cases the largest savings was from the first to the second stimulus, i.e. as soon as one stimulus had been habituated to, subsequent stimuli were habituated to much more rapidly. Further savings occurred with the third and fourth stimuli presented, until a mean of only one stimulus presentation was needed before habituation criterion was reached. The number of animals requiring zero trials

to reach criterion is shown in Table IV. These results show that whilst habituation was specific to stimulus frequency, nonetheless there was considerable savings in habituation over a series of stimuli.

TABLE IV

Number of rats taking zero trials to reach habituation criterion

Stimulus position	Stimulus			
	7 kHz	9 kHz	C	R
1	0	0	0	0
2	4	3	3	3
3	4	4	5	4
4	5	5	5	5

Discussion

About 75% of the rats showed habituation to be specific to both the stimulus frequency and melody. In making such an interpretation it is essential that the stimuli do not differ in any other parameters, such as subjective intensity. This was controlled for in this study by the use of stimulus frequencies on the flat portion of the rat's audiometry curve (Clack and Harris, 1963; Gourevitch and Hack, 1966). It was also confirmed empirically by the fact that all the stimuli used were found to be equally distracting. This factor does not seem to have been controlled for in the Miller and Murray study.

The savings in habituation when a series of tone stimuli was presented emphasizes the importance of counterbalancing the order of stimulus presentation when testing for the specificity of habituation. By the time the fourth stimulus was presented habituation was complete in an average of one trial.

The more rapid habituation with successive stimulus presentations could have been a function of the number of previous trials, regardless of the number of *different* stimuli presented. In which case, there would be a negative correlation between the trials to habituate to the second stimulus and the number of previous trials received. Pearson's Product-Moment correlation coefficient was calculated and was $+0.10$, which was not significant ($P > 0.10$). Thus, the savings in habituation to the second stimulus was not simply a function of the time spent in the situation, i.e. the number of trials.

A more important alternative remains. The savings in habituation may have been a function of the number of preceding different stimuli, regardless of their similarity to each other. Coombs (1938) found no savings when a series of widely differing auditory stimuli was presented. Porter (1938) investigated savings in GSR habituation to a series of inter-modal stimuli, but whilst no significant differences were found he claimed that some savings in habituation had occurred. In order to investigate this a second experiment was designed in which a series of inter-modal stimuli was presented.

Experiment II

Apparatus

The apparatus was the same as that described in the previous experiment. In addition, a 24-V, 2·8-W light bulb was placed above the water spout. The light intensity at the

spout was 0·5 lm/ft² (1 ft² = 0·0929 m²). The tone intensity was 77·5 dB and only one tone frequency, 9 kHz, was used. Both the tone and light stimulus presentations were 20 s long, and only one presentation per day was given.

Subjects

Forty-eight male hooded rats were taken from the colony stock at weights of 300–350 g, and housed and deprived as in the previous experiment.

Procedure

Pre-habituation training sessions were given, as before, for five days. The rats were then randomly assigned to four groups, A, B, C and D, with 12 rats in each group.

Group A rats received daily presentations of a light stimulus until they reached habituation criterion, whereupon a tone stimulus was presented until criterion was again reached. Group C received a tone stimulus first, and then a light. Groups B and D acted as matched controls for Groups A and C respectively. They received no stimulus presentations of any kind until, in each case, the rat to which each was matched habituated to the first stimulus. At this point, Group B rats received tone presentations, and Group D rats light presentations, until habituation criterion was reached.

The measures of distraction and habituation, and the assessment of criterion, were the same as in the previous experiment.

TABLE V

Savings in habituation to inter-modal stimuli

Group		1st stimulus	2nd stimulus
A		Light	Tone
	Distraction ratio	0·89	0·67
	Trials to criterion	11·2	5·4
B			Tone
	Distraction ratio		0·65
	Trials to criterion		5·7
C		Tone	Light
	Distraction ratio	0·68	0·92
	Trials to criterion	5·8	11·5
D			Light
	Distraction ratio		0·84
	Trials to criterion		10·5

Results

Table V shows the initial distraction to the first presentation of each stimulus, and the trials to reach habituation criterion. Comparing the distraction to the first stimulus in each series it is immediately apparent that the light was more distracting than the tone (a ratio of 0·89 compared with 0·68). Almost twice as many trials were needed to habituate to the light than to the tone (11·2 and 5·8). Since the stimuli were not equally distracting interpretations about savings in habituation must be made with care, and only by comparison with the relevant control group scores.

If there was savings in habituation, Group A should have habituated to the tone more rapidly than Group B, because of prior habituation to the light. The groups took 5·4 and 5·7 trials respectively to habituate, which is not a significant difference. Similarly, a comparison between Groups C and D should reflect any savings in habituation to the light as a result of prior habituation to a tone. The two groups took, respectively, 11·5 and 10·5 trials to habituate, which is not a significant difference. From Table V it can also be seen that the distraction to the second stimulus was unaffected by prior habituation to the first (comparing Groups A and B, and C and D).

The experience of a number of blank trials prior to the first introduction of a stimulus could have affected the initial distraction and subsequent habituation to a stimulus. A comparison of the responses of Groups A and D to the light stimulus shows that the blank trials had no effect (distraction ratios of 0·89 and 0·92 respectively, and 11·2 and 11·5 trials to criterion). Similarly, a comparison of the responses of Groups C and B to the tone reveal that the blank trials had no effect.

Conclusions

The results of the first experiment indicated that habituation was specific to the frequency of simple auditory stimuli, as well as to the melody of patterned stimuli. These results confirm, at a behavioural level, the previous findings from electrophysiological measures of habituation. Since only one stimulus presentation per day was given, the acquisition conditions were such as to preclude any short term physiological processes, e.g. fatigue, underlying the behavioural change. This, together with the specificity of the changes, would corroborate suggestions that habituation be regarded as a form of learning.

In spite of the specificity of habituation, there was also considerable savings in habituation when a series of similar (e.g. tones, differing only in frequency) stimuli was presented. This savings was not a function of the number of trials in the situation, nor of the number of preceding stimuli *per se*, as there was no savings if the stimuli were very different from each other (e.g. of different modalities).

The greatest savings occurred between the first and second stimulus in Experiment I, and hence any savings should have been apparent in Experiment II, although only two stimuli were presented. These results suggest that savings in habituation was a phenomenon more related to stimulus generalization than to a generalized drop in responsivity as a result of exposure to a series of stimuli.

The savings in habituation when a series of similar stimuli is presented has important implications for the methodology employed in assessing the specificity of habituation. The test stimuli must be presented in counterbalanced order, and the use of more than two test stimuli would make interpretations difficult, due to the large amount of savings.

References

ABEL, E. (1971). Habituation as a factor in early handling. *Journal of Comparative and Physiological Psychology*, **74**, 219–21.

BARRETT, R. J. and OAKLEY, S. R. (1969). Attenuation of habituation by electro-convulsive shock. *Journal of Comparative and Physiological Psychology*, **69**, 133–5.

BISHOP, P. D. and KIMMEL, H. D. (1969). Retention of habituation and conditioning. *Journal of Experimental Psychology*, **81**, 317–21.

BLANCHARD, R. J., SHELTON, V. F. and BLANCHARD, D. C. (1970). Historical effects of stimulus exposure: Readiness to eat and object exploration. *Learning and Motivation*, **1**, 432–44.

CLACK, T. and HARRIS, J. (1963). Auditory thresholds in the rat by a two-lever technique. *Journal of Auditory Research*, **3**, 53–63.

COOMBS, C. (1938). Adaptation of the galvanic response to auditory stimuli. *Journal of Experimental Psychology*, **22**, 244–68.

CORMAN, C. (1967). Stimulus generalisation of habituation of the GSR. *Journal of Experimental Psychology*, **74**, 236–40.

FILE, S. E. (1969). Habituation—a parametric study of the behavioural phenomenon. Unpublished University of London Ph.D. thesis.

FOKKEMA, S. (1966). Orientation reflex conditioning and generalisation phenomena. *Nederlandsch tijdschrift voor psychologie en haar grensgebieden*, **21**, 82–101.

GOUREVITCH, G. and Hack, M. (1966). Audibility in the rat. *Journal of Comparative and Physiological Psychology*, **62**, 289–91.

HARRIS, J. (1943). Habituatory response decrement. *Psychological Bulletin*, **40**, 385–422.

HORN, G. (1967). Neuronal mechanisms of habituation. *Nature*, **215**, 707–11.

KIMMEL, H. and GOLDSTEIN, A. (1967). Retention of habituation of the GSR to visual and auditory stimulation. *Journal of Experimental Psychology*, **73**, 401–4.

MEYERS, W. and GULLIKSON, A. (1968). The evoked heart-rate response: the influence of auditory stimulus repetition, pattern reversal and autonomic arousal level. *Psychophysiology*, **4**, 56–66.

MILLER, N. E. (1967). Certain facts of learning relevant to the search for its physical basis. In QUARTON, G., MELNECHUK, T. and SCHMITT, F. (Eds), *The Neurosciences: A Study Program*. New York: Rockefeller University Press.

MILLER, J. and MURRAY, F. (1966). The guinea-pig's immobility response to sound: threshold and habituation. *Journal of Comparative and Physiological Psychology*, **61**, 227–33.

PORTER, J. (1938). Adaptation of the GSR. *Journal of Experimental Psychology*, **23**, 553–7.

SHARPLESS, S. and JASPER, H. (1956). Habituation of the arousal reaction. *Brain*, **79**, 655–81.

SOKOLOV, E. (1960). Neuronal models and the orienting reflex. In BRAZIER, M. (Ed.), *CNS and Behaviour*, **2**, 187–276.

STELLAR, E. and HILL, J. (1952). The rat's rate of drinking as a function of water deprivation. *Journal of Comparative and Physiological Psychology*, **45**, 96–102.

WEISKRANTZ, L. (1968). Emotion. In WEISKRANTZ, L. (Ed.), *Analysis of Behavioural Change*. New York: Harper & Row.

WICKELGREN, B. (1967). Habituation of spinal motoneurons. *Journal of Neurophysiology*, **30**, 1404–23.

ZIMNY, G. and SCHWABB, L. (1965). Stimulus change and habituation of orienting responses. *Psychophysiology*, **3**, 103–15.

THE DISRUPTIVE EFFECT OF AN IRRELEVANT MESSAGE IN DICHOTIC LISTENING TO WORD-LISTS

C. M. HOLLOWAY

Medical Research Council, Applied Psychology Unit, Cambridge

The results of previous experiments suggest that when listeners are asked to repeat back prose arriving at one ear, their performance is unaffected by the presence (in the other ear) of another independent prose passage in the same voice and of equal intensity. In the present study dichotic messages were employed in which the words were presented quasi-simultaneously and were not connected in meaning. Considerable performance decrement was found from a message on the irrelevant ear even if that message was less intense than that on the relevant ear. Different kinds of error occur, but the principle form of error is a response which is incorrect but nevertheless clearly related to the word presented to the relevant ear. The results extend the findings of Treisman (1967, 1970) who found considerable interference in dichotic listening to pairs of words exactly matched for time of onset; and suggest a general difficulty of selective attention when non-prose material is used.

Introduction

The results of early experiments on dichotic listening suggest that when listeners are asked to repeat back or "shadow" a prose passage arriving at one ear, their performance is unaffected by the presence of another message by the same voice and of equal intensity in the other ear (Cherry, 1953). Only when the irrelevant message is considerably louder than the relevant message (Egan, Carterette and Thwing, 1954; Moray, 1958) or when the listener must reject more than one irrelevant message (Treisman, 1964) is there a decrement in performance. These latter results have been taken as a reflection of the efficiency with which the separation of dichotically presented messages can "normally" be achieved. However, it is possible that the efficiency of message selection may be overestimated by experiments in which a prose shadowing technique is utilized.

There is a great deal of informal evidence to the effect that listeners who are asked to shadow prose passages tend to group their responses into phrase-lengths. The delay in response in such cases is considerably greater than it potentially might be, as judged by response times to individually presented words (cf. Pollack, 1963). It is thus not implausible to suggest that the listener normally monitors the output from a word recognition system in order to assemble meaningful phrases before overt responding. Given such a monitoring function and the general redundancy

of prose passages the listener may respond accurately to a given prose message even if individual words are not correctly recognized due to some (as yet unspecified) source of difficulty in rejecting the message on the unwanted ear. Performance would remain optimal because the listener's knowledge of the language in general and the "sense" of the prose passage in particular would allow him both to detect any anomalous output from the word recognition system, and also to replace such an output with a contextually appropriate word (which more often than not would be correct).

The operation of a monitoring function in prose shadowing experiments might also be expected to affect the type of errors obtained for those dichotic listening conditions which are sufficiently adverse to produce some level of performance decrement. For example, errors attributable to the reporting of words from the message on the irrelevant ear seem to be atypical. None are reported by either Egan *et al.* (1954) or Moray (1958) for loud irrelevant messages. Treisman (1964) using two irrelevant messages, reports that few errors of intrusion occurred. However, the presence of a monitoring function would be particularly likely to effect the elimination of "intrusions" from overt response.

If random sequences of words are used it should be possible to eliminate the operation of any monitoring function and obtain a more realistic estimate of the extent and nature of the difficulty of dichotic listening. A recent experiment by Treisman (1967) represents such an experimental paradigm. She presented a message to each ear in the same voice and at the same intensity. Each message consisted of a string of random words, and the words at each ear were matched for time of onset to within 1 ms. [Full experimental details of this experiment can be found in Treisman and Riley (1969).] With this task there was a high error rate associated particularly with the occurrence of responses corresponding to words from the irrelevant message. Although Treisman ascribed the general difficulty of her task to the exact matching of the time of onset of the words on each ear it might be suggested that her results are in part due to the use of random word sequences *per se*, irrespective of the way in which the words in each ear are matched for time of onset.

In the present experiment random word sequences are used but the words at each ear are not exactly matched for time of onset. The loudness of the irrelevant message is varied systematically. Using this paradigm it should be possible to judge the extent to which Treisman's results represented a special case due to the matching of onset times rather than an example of the general difficulty of dichotic listening in the absence of message redundancy.

Method

Apparatus and Stimulus Material

Word lists were presented simultaneously, one to each ear, via Canada HA10 headphones driven by a Vortexion two-channel tape recorder. A simulated telephone link employing pulse modulation coding was introduced between recorder and headphones. This allowed the independent manipulation of both instantaneous signal-to-noise ratio and absolute intensity on each channel. Each word list consisted of 12 pairs of monosyllables drawn at random from a set of 500 in the revised CNC lists for auditory testing (Peterson and Lehiste,

1962). The words on the ear to be attended were preceded by a count of "three, two, one". The words were uttered at a rate of 88 per min and the words on the two channels were synchronized within the limits of the experimenter's perceptual-motor-skill. The mean lag between words on the two channels was 8 ms, attended channel leading, and the standard deviation was 60 ms.

Throughout the experiment the relevant message was presented to the preferred ear at 75 dBA (peak value). The intensity of the irrelevant message in relation to that of the relevant message was varied in five steps: +20 dB, +10 dB, +0 dB, −10 dB and −20 dB.

To manipulate overall task difficulty four signal-to-noise ratios were used (+0 dB, +4 dB, +8 dB, and +16 dB), both messages carrying the same ratio. The noise at each ear was uncorrelated. At more unfavourable signal-to-noise ratios response times are slower, so that manipulation of signal-to-noise ratio is similar to manipulation of rate of presentation in controlling task difficulty (Holloway, 1970).

The signal-to-noise ratio of the messages and the relative intensity of the two messages were varied independently to give 20 conditions.

Subjects and Procedure

Ten subjects were drawn from the Applied Psychology Unit subject panel. These subjects were female, with a mean age of 40 and a range of 21–60 years. The only restriction on selection was that no subject reported any noticeable hearing difficulty. Each subject was presented with 200 pairs of 12-item word lists in the course of a 2-h testing session (with three 5-min rest periods). After 15-min practice at repeating back, and familiarization with the experimental conditions, each of the five relative intensities of the irrelevant message were presented five times each in a latin square design. At each relative intensity each of the signal-to-noise ratio conditions were given in random order.

The subjects were required to repeat back as many words as possible, as accurately as possible, from the relevant message. They were specifically instructed to attempt to keep up with the rate of presentation and not to lag behind the words as they were presented. "Storage" of groups of words was actively discouraged. Responses were tape-recorded on good quality equipment for later analysis.

Results

In the analysis of results for all experiments reported in this paper, the data have normally been subjected to analysis of variance, treating subjects as replications. Unless otherwise stated, it may be assumed that data have been given an arc-sign transformation to provide a best fit to the normality and homogeneity of variance assumptions. The criterion of significance of F ratios is the 0·01 level and non-significance is assumed to imply a probability value of less than 0·05, unless otherwise stated.

Items Correct, Relevant Messages

In the first analysis responses were scored simply in terms of the percentage of items from the total set of words presented which were correctly reported from the relevant message (Fig. 1).

An analysis of variance shows a significant main effect of both signal-to-noise ratio of the messages and the relative intensity of the irrelevant message. Each level is significantly different from each other level within each of these two variables ($P < 0·01$, Scheffe test). The interaction between these two variables is not significant.

Items Intruding from the Irrelevant Message

The second analysis concerns intrusion errors, which were defined as intrusions of words from the irrelevant message, or the occurrence of responses "rhyming" with the word in the irrelevant message, but not "rhyming" with the word in the relevant message.

A *rhyming error* is defined as any response which has the same vowel sound as a particular word, or has identical initial and final consonant clusters. Such responses are a common form of misperception in a speech recognition test. Examples are shown in Table I. The rhyming error is thus a crude index of the failure to process adequately a particular word for responses nevertheless clearly associated with that word in some way.
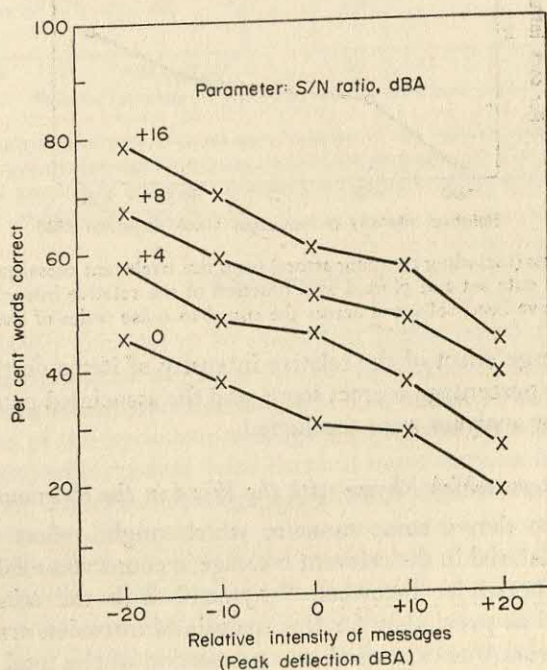
FIGURE 1. The percentage of words correct from the relevant message is plotted as a function of the relative intensity of the irrelevant message. The parameter is the signal-to-noise ratio of the messages.

TABLE I

Rhyming Errors

Stimulus	Rhyming errors		Non-rhyming errors	
	Common vowel	Initial *and* final consonant clusters common	Initial *or* final consonant clusters common	No common elements
SAM	FAT, PAN	SUM, SEEM	SIP, LOOM	TIP

Intrusions and "rhyming" responses associated with the irrelevant message were summed and expressed as a percentage of the total data set; few intrusions were found to occur (Fig. 2). Even with the most intense value of the irrelevant message intrusions did not account for more than 2% of the data set. The paucity of intrusions prevented any sensitive statistical test of the intrusion scores. However, by pooling the data for the 10 subjects, it was possible to show a significant main effect of the relative intensity of the irrelevant message.
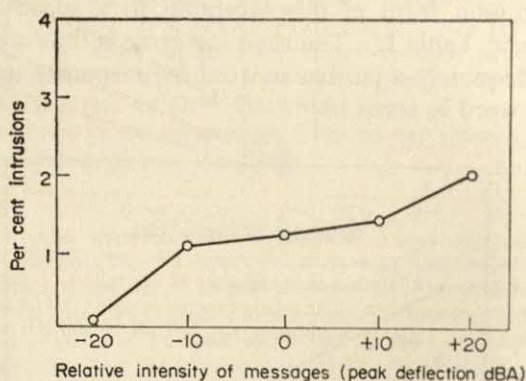
FIGURE 2. Intrusions (including rhyming errors) from the irrelevant message are expressed as a percentage of the total data set and plotted as a function of the relative intensity of the irrelevant message. The data have been collapsed across the signal-to-noise ratios of the messages.

In view of the large effect of the relative intensity of items on the irrelevant ear as measured by the percentage correct score, and the associated paucity of intrusion errors, further error analyses were conducted.

Incorrect Items which Rhyme with the Word in the Relevant Message

In an attempt to derive some measure which might reflect the accuracy of processing of the material in the relevant message, a count was made of the number of responses which, whilst incorrect, "rhymed" with the stimulus word. A "rhyme" is defined as previously for the analysis of intrusion errors (q.v.). The number of such errors was expressed as a proportion of the total set of responses uniquely attributable to the relevant message, i.e., correct and "rhyming" responses from the relevant message. The pattern of results is shown in Figure 3(a).

An analysis of variance indicates that the proportion of "rhyming" errors rose significantly as the relative intensity of the irrelevant message was increased. Each level of relative intensity gave a significantly different level of performance from each other level ($P < 0.01$, Scheffe test). The signal-to-noise ratio affected overall performance, but did not interact with relative intensity.

Omission Responses

Omissions were expressed as a percentage of the total data set. An analysis of variance confirms the impression given in Figure 3(b) that omissions vary significantly as a function of both signal-to-noise ratio and the relative intensity of the non-relevant message, with no interaction between the two variables.
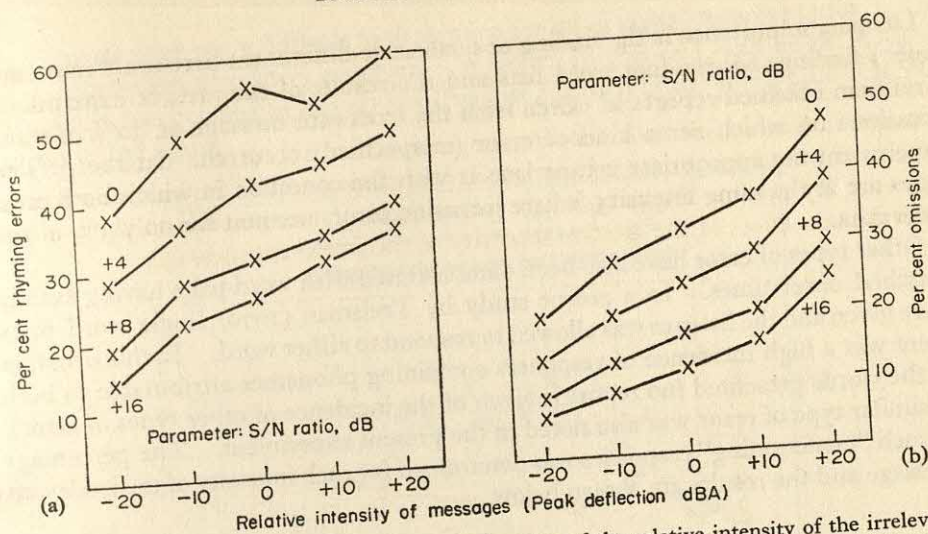
FIGURE 3. An analysis of errors, plotted as a function of the relative intensity of the irrelevant message. The parameter is the signal-to-noise ratio of the messages. (a) The percentage of rhyming errors from the total number of responses uniquely attributable to the relevant message. (b) The percentage of omissions from the total data set.

Discussion

Progressively larger decrements in shadowing performance are elicited by progressively greater intensities of the irrelevant message. This trend is apparent for all measures of performance decrement. The extent of the decrement at low relative intensities of the irrelevant message is in accord with the results obtained by Treisman (1967) with random word lists but more extreme in that a decrement is produced by an irrelevant message *less* intense than the relevant message, a finding not previously demonstrated. Because of the wide distribution of onset times between different word pairs in the present experiment the results allow an important generalization of Treisman's findings with word pairs matched for the time of onset to within 1 ms and it seems reasonable to conclude that prose shadowing experiments have seriously underestimated the difficulty of the dichotic listening task.

It has been argued that the success of prose shadowing in the dichotic listening situation might be attributable to the listener's opportunity to make good any confusions or misperceptions at the word level by using the general sense of the message being followed. It is difficult to quantify how successful such a stratagem might be in eliminating erroneous responses, and other differences between the two types of test material may in some measure account for the observed differences in performance. The degree of overlap of words in two prose messages may well be less than is found in the present experiment, and the prosodic pattern of the prose more predictable. Without recourse to further experimentation it is only possible to say that message redundancy, in the form of either semantic or prosodic constraints, plays an important role in the success of selective attention in dichotic listening.

Of some importance is the finding of qualitative differences between Treisman's (1967) findings for random word lists and the results of the present experiment. Treisman obtained reports of words from the irrelevant message on 50% of those occasions on which some kind of error (unspecified) occurred. In the present experiment the appropriate comparison is with the condition in which both messages are at the same intensity, where intrusion errors account for only 3% of the error data.

Other types of error have also been demonstrated with word pairs having exactly matched onset times. In a recent study by Treisman (1970) single word pairs were given and the listener was allowed to respond to either word. In this situation there was a high incidence of responses containing phonemes attributable to both of the words presented (no report is given of the incidence of other types of error). A similar type of error was also noted in the present experiment. The percentage of such "confounding" responses was determined for each intensity of the irrelevant message and the results are shown below.

TABLE II

Percentage of "confounding" responses out of all errors of commission

Irrelevant/relevant speech level (peak deflection dBA)				
−20	−10	0	+10	+20
6·1	11·8	15·2	16·6	14·3

Although Treisman does not represent confounding responses as a proportion of the total set of errors, it appears that they form a larger percentage of the error total in her experiment than they do in the present experiment.

In general, a comparison of the present data with Treisman's results shows that whilst similar kinds of errors occur, their size as a proportion of the error total differs. It is tempting to explain such differences in terms of the degree of matching of onset times between the different experiments. The type of error most frequently found with word pairs having exactly matched onset times (intrusions and confounding responses) might then be viewed as a class of errors superimposed on a general performance decrement more typically described in terms of the error pattern established in the present experiment. In order to evaluate such a hypothesis the data from the present experiment have been reclassified in terms of the relative lag between onset times of individual word pairs.

The relative onset time of words was estimated by slowly and repeatedly passing a word from the relevant message, at full gain, over the playback head of the tape recorder to locate the points of onset. After marking the tape the appropriate word from the irrelevant message was then located in a similar manner, and marked. The distance between marks on the magnetic tape was then used to estimate the time lag between onset to an accuracy of about ±5 ms. Only those words recorded at the most favourable signal-to-noise ratio were processed in the manner described,

since onsets were more difficult to determine for other signal-to-noise ratios. The total distribution of relative onset times is shown in Figure 4, and on the basis of this plot it was decided to inspect performance for each of the seven class intervals shown in the figure.
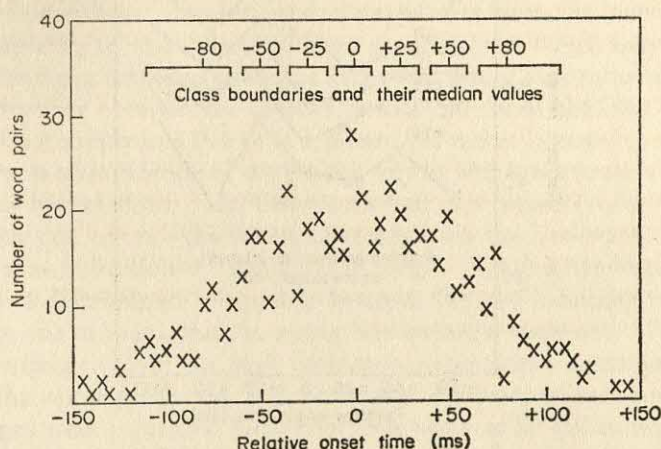
FIGURE 4. The number of word pairs is plotted as a function of the relative onset times of words in a pair. Positive values represent instances in which the word which the listener is instructed to attend is leading in time. The class boundaries and associated median values used in the analysis of the effects of relative onset time are also shown.

In the first analysis a percentage correct score was used as the most general indicator of performance decrement. The data are shown in Figure 5. An analysis of variance, treating subjects as replications and using an arc-sign transformation, shows a significant main effect ($P < 0.01$) for relative intensity of the irrelevant message, but no significant effect for relative onset time and no significant interaction between relative intensity and relative onset time.

The analysis of performance decrement in terms of a percentage correct score could conceal other changes in the relative frequencies of particular kinds of error at different relative onset times, but this kind of effect is not apparent in the present case. Figure 6 shows the frequency of different kinds of error plotted as a function of relative onset times. In view of the small amount of data on which some of these error scores are based, the data has been collapsed across the five relative intensities of the irrelevant message. A one-way analysis of variance was applied to each of the different kinds of error score but no significant effect of relative onset time was obtained.

Since individual word pairs are strung together in sequences, the present results would be expected to produce an attenuated effect of relative onset time due to the carry-over of difficulty with one word pair to the succeeding word pair. However, there is absolutely no tendency at all for the particularly well matched word pairs to provide an especial form of difficulty in this experiment. Of particular importance in this respect is the uniformity of the balance of different kinds of error across the range of relative onset times, and the discrepancy between the pattern of errors observed here and that observed by Treisman. The present experiment thus
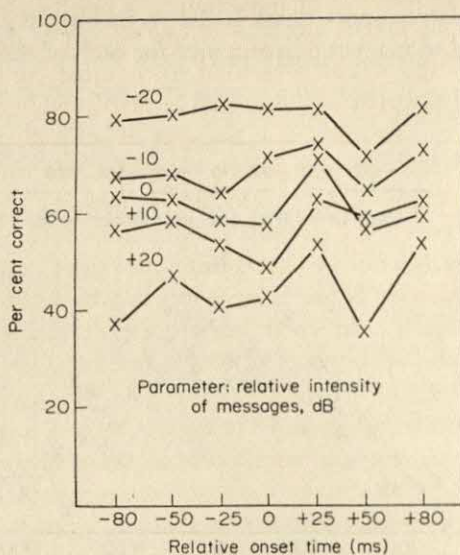
FIGURE 5. The percentage of words correct from the relevant message is plotted as a function of the relative onset times of words in a pair. The parameter is the relative intensity of the irrelevant message. Only the data for words presented at the most favourable signal-to-noise ratio are shown.
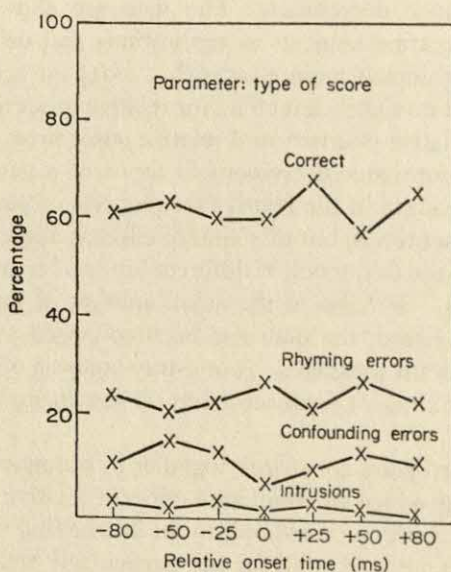
FIGURE 6. Errors are plotted as a function of relative onset times of words in a pair. The individual curves represent the different error scores. The data have been collapsed across the range of values of relative intensity of the irrelevant message. Only the data for words presented at the most favourable signal-to-noise ratio are shown.

serves to demonstrate an effect of dichotic listening performance of the elimination of message redundancy, which is not necessarily related to the degree of matching of time of utterance of the words presented to each ear, and differences between Treisman's results and the present ones must be attributable to other differences in experimental procedure. In particular, the type of subject, the instructions given, and the complexity of experimental design are all factors which may have elicited different listening or response strategies in the two sets of experiments.

The discrepancy between the present results and those obtained by Treisman is of interest not simply as a failure to replicate but rather because the error pattern may be taken as an indication of the precise way in which the process of selectively attending has broken down. In Treisman's first experiment (1967) the high incidence of intrusion errors—the report of words from the irrelevant message—was interpreted as an indication of "channel switching". On this hypothesis the listener is supposed to discriminate adequately between the two incoming messages, and fully process one of them, but the wrong one on some occasions. In Treisman's second experiment (1970) the high incidence of responses containing phonemes from both the words presented, and the similar incidence of such responses when both messages were presented binaurally, was taken as an indication of "channel fusion". On this hypothesis it is supposed that the listener fails to make use of cues to the ear of origin of the messages. Treisman's two hypotheses may be seen as pointing the antithesis between *wrongly* selecting and *not* selecting as explanations of failure to attend selectively to the target message. Although these hypotheses deserve greater elaboration than it is appropriate to give here, it suffices to note that whilst support has been obtained for *both* these apparently mutually exclusive hypotheses in two different experiments by Treisman, *neither* gains support from the present data. Thus, the channel switching hypothesis fails to account for the fact that, at any of the relative intensities of the irrelevant message, the number of intrusion errors is very small. The channel fusion hypothesis also fails on the grounds of the paucity of intrusion errors, but specifically at those intensities of the irrelevant message greater than the intensity of the relevant message. If the words from both ears are treated as if not differentiated by ear of origin, the louder of the two words would be expected to be reported, as happens when two words are presented to the same ear.

The present experiment points to the necessity of conducting a detailed error analysis in any investigation into the causes of performance decrement in dichotic listening, and the need for a more detailed specification of those conditions under which different kinds of error can be elicited. Only in this way will it be possible to specify the limitations of the selective attention process which must be invoked in order to account for the performance decrement elicited in those experiments employing non-redundant word lists.

References

CHERRY, E. C. (1953). Some experiments on the recognition of speech with one and with two ears. *Journal of the Acoustical Society of America*, 25, 975–9.

EGAN, J. P., CARTERETTE, E. C. and THWING, E. J. (1954). Some factors affecting multi-channel listening. *Journal of the Acoustical Society of America*, 26, 774–82.

HOLLOWAY, C. M. (1970).   Paced recognition of words masked in white noise.   *Journal of the Acoustical Society of America*, **47**, 1617–8.

MORAY, N. (1958).   The effect of relative intensities of messages in dichotic shadowing. *Language and Speech*, **1**, 110–3.

PETERSON, G. E. and LEHISTE, I. (1962).   Revised CNC lists for auditory testing.   *Journal of Speech and Hearing Disorders*, **27**, 62–70.

POLLACK, I. (1963).   Reaction times to unknown word sets in Noise.   *Language and Speech*, **6**, 189–95.

TREISMAN, A. (1964).   The effect of irrelevant material on the efficiency of selective listening.   *American Journal of Psychology*, **77**, 533–46.

TREISMAN, A. (1967).   Reply to "Comments on Selective Attention; Perception or Response".   *Quarterly Journal of Experimental Psychology*, **19**, 365–7.

TREISMAN, A. and RILEY, J. G. A. (1969).   Is selective attention selective perception or selective response? A further test.   *Journal of Experimental Psychology*, **79**, 27–34.

TREISMAN, A. (1970).   Perception and recall of simultaneous speech stimuli.   *Acta Psychologica*, **33**, 132–48.

SPATIAL AND VERBAL CODES AND THE CAPACITY OF STM†

LESLIE HENDERSON‡

Department of Psychology, University of Waterloo, Ontario

Subjects were required to report an array of consonants presented tachistoscopically for various durations. Concurrently, they held in memory a load of auditory digits, a load of spatial locations or no load. These several forms of information did not seem to compete for a fixed capacity short term store. Consonant report was unaffected by carrying a spatial load and barely affected by carrying the auditory digit load. An auditory–verbal loop cannot play a mandatory role in the coding and storage of visual information. Some criticisms of structural interpretations of immediate memory are suggested.

Introduction

One interpretation of the span of immediate memory is that it is a structural limit within which coding options must operate (e.g. Miller, 1956). Neisser's (1967) description of "slot" theory is an embodiment of this view. An alternative interpretation, which might be described as functional, holds that the span is no more than an indication of the encoding opportunities available in a task characterized by rapid presentations of items with little semantic or syntactic structure. These coding dimensions are usually restricted to physical characteristics of the item such as their sound (Coltheart & Geffen, 1970) or to temporal features such as their recency. It is possible that the restricted coding dimensions in the immediate recall situation may prevent an adequate indexing of material for retrieval.

The structuralist interpretation of the span may be defended by defining immediate recall so as to exclude supra-span situations. For example, supra-span performance may be attributed to the transfer of material to a secondary store. Such a manoeuvre, however, is either tautologous or is reducible to the functionalist interpretation translated into memory-box language.

Sanders and Schroots (1969) report an experiment in which a list of six auditory consonants is supplemented by one through six additional items; these items, according to the condition, are further consonants, digits, tones or spatial locations. The composite list is presented for immediate recall. The results in terms of total number of items recalled, show increasing span as the similarity of the supplementary category decreases. The effect approaches additivity of memory spans when consonants are followed by spatial locations. The obvious possibility

---

that the consonants are coded acoustically and the locations spatially does not escape the authors.

This paradigm has recently been taken a step further by Scarborough (1971), who has shown that when an auditory sequence of eight digits is followed by a tachistoscopic display of digits or letters and recall is post-cued for one of the modalities of presentation, visual performance is unimpaired by the presence of the auditory load. This strongly suggests that retention of the visual display did not rely upon auditory STM, especially in the case where both input sets were digits. Even when the visual display is followed by a mask and the cue is delayed by 2 s the auditory load only imposes a 12% decline on visual report performance. Scarborough argues on the basis of these results that subjects have a post-iconic storage capacity of 4–5 items without resort to implicit speech. If the duration of the display permits the subject supplements this capacity by naming items. This he can only do when no auditory load is carried concurrently. The prediction of a greater detrimental effect of auditory load with longer visual displays is borne out by the data.

It is tempting to identify Scarborough's non-verbal store with a post-iconic visual code. This temptation is increased by finding that the capacity for storing spatial locations presented seriatim in a discrete matrix has been reported by Sanders (1968) to be 3–4. (This is essentially the visual task studies by Sanders and Schroots (1969).) Sanders (1968) also reports that this spatial memory unlike verbal memory, shows no recency effect, is insensitive to rate of presentation, exhibits visual confusions and is sensitive to the set of alternatives.

Posner has argued for a similar visual code on the basis of a quite different series of experiments in which subjects classify two letters as the same (in name) more rapidly when the letters are visually identical than when they share the same name but vary in case (Posner, Boies, Eichelman and Taylor, 1969). Posner (1969) has discussed the properties of this visual code and argued for its distinction from iconic memory on such grounds as its apparent capacity limit of 3–4 items and its greater persistence.

The new magic number 4 ± 1 which emerges from the diverse experiments of Sanders, Scarborough and Posner coincides with the span of apprehension for tachistoscopic report of items *in their correct position* which Sperling (1960; 1963) has attributed to the limits of immediate verbal (auditory) memory. As we have seen, this cannot be so. These experiments are quite fatal for theories which hold that visual information survives the icon only in auditory–verbal form (e.g. Sperling, 1967; Glanzer and Clark, 1963; Atkinson and Schiffrin, 1969).

The simple expedient of postulating parallel visual and verbal codes seems to provide a better approximation to the data. (See also Henderson, 1972.) However, it does not account for the continuous increase in memory span obtained by Sanders & Schroots (1969) using auditory material of increasing categorical diversity. Nor does it give a satisfactory account of the following experiment.

Stimulus Material

The tachistoscopically presented material comprised arrays of nine consonants drawn from two alphabets. The consonants were Futura Bold reversed Letraset transferred to

white cards, yielding a contrast ratio of about 1 : 10. They were arranged in a 3 × 3 matrix. At the viewing distance of 1,250 mm each consonant subtended an angular height of 0·25° and the entire matrix subtended 1·5° in both dimensions.

The arrays were drawn randomly without replacement from one of two alphabets. The AS (acoustically similar) alphabet consisted of the 9 consonants DPCTGVZBW : eight of these share the redundant phoneme ē (if Z is pronounced "zee") and W has the most phonologically complex name. The AN (acoustically neutral) alphabet comprised the 20 consonants with Y treated as a vowel. In a given condition an array was never used twice, AN and AS arrays were randomly mixed but the ratio of AN to AS arrays was always 2 : 1.

A visual mask was constructed from the same letter material using E and I as components. This mask met the requirements that it rendered the letters illegible when superimposed optically (fulfilling the desideratum of the integration theory of masking) and that it utilized components of the same spatial frequency as the stimulus (Henderson, Coles, Manheim, Muirhead and Psutka, 1971). A grey fixation point on a white card constituted the third field.

The material for the spatial task consisted of a set of small cards each of which contained a 5 × 5 matrix of cells six of which were coloured red. The cards were chosen from a larger set with the requirement that the disposition of the filled cells appear as random and uncodable as possible. A set of empty matrices were used for the responses.

For the auditory task series of nine digits were recorded on tape. All series consisted of the digits 1–9 in scrambled order. They were spoken at the rate of 1/s with a slight grouping in threes.

Response blanks for the Main Task (tachistoscopic consonants) and the Auditory Task (digits) were constructed. Those for the Main Task contained an empty 3 × 3 matrix.

Procedure

Subjects performed four sessions of about an hour each and a post-experimental interview. The first three sessions comprised the original design. Each of these sessions was devoted to one of the three Load Conditions, No Load (Main Task only), Auditory Load, and Spatial Load. The order of conditions was balanced across subjects so that each condition occurred equally often in each serial position.

Within each session three blocks of trials were run on the Main Task. These blocks were at an exposure duration of 100, 400 and 1,000 ms. The order of exposure durations used (ascending or descending) was varied orthogonally to the order of Load Conditions. A block comprised 24 trials, 8 AS trials and 16 AN.

The Main Task displays were presented by a Scientific Prototype Model GB three-field tachistoscope. When the subject pressed the trigger the fixation-adapting field was replaced by the consonant array for the appropriate exposure duration and this in turn was replaced by the mask for 1 s. Then the adapting field returned. The mask was aligned so as to superimpose on the stimulus array.

Immediately after a stimulus presentation the subject recorded his responses on the blank matrix. He was instructed to fill all the squares, attempting to reproduce the items in their correct locations and guessing when necessary.

In the spatial load condition the subject was shown a random pattern for 4 s. This was followed immediately by performance on the main task. After Main Task recall the subject attempted to reproduce the random pattern by making six crosses in a blank matrix.

In the Auditory Load condition the digit series was presented immediately before Main Task performance. Forced choice recall was made after recall for the Main Task.

After the data from the first three sessions had been part analysed, it seemed appropriate to have data on performance at a higher level. Accordingly a fourth session was arranged. In this session exposure duration of 1250 ms was used in the No Load and Auditory Load conditions. Data were also collected in this session on performance in the Auditory Task without the Main Task interpolated.

Subjects

Six subjects performed the experiment, three of either sex. The subjects were undergraduates and graduates in the Psychology programme at the University of Waterloo. They were paid $2.00 per session for their participation.

Results and Discussion

The Main Task responses were scored for items correct in the correct location. These scores were computed separately for AS and AN stimuli. The same scoring procedure was applied to the responses on the Auditory and Spatial Tasks. This scoring procedure is especially conservative in the Spatial Task as subjects often transpose a pattern by one row in their responses. When such responses are scored by a correctly-identified-cell criterion they may achieve a score no better than chance.

Main Task performance was also scored by the more lenient criterion of items correct regardless of position. In this case credit was only given once for a given consonant regardless of how often it occurred in the response. In fact responses were seldom repeated in this way.

TABLE I

Average number of consonants reported in the correct position in the Main Task

Subject	Acoustically similar				Exposure (ms)	Acoustically neutral			
	100	400	1000	1250		100	400	1000	1250
S.S.	2·25	3·1	2·6	3·9	No Load	1·4	3·0	3·2	4·0
	1·9	3·0	3·0	3·0	Auditory Load	1·8	2·75	3·7	4·1
	2·0	3·4	3·1		Spatial Load	2·5	2·9	4·5	
AB	3·4	4·9	5·75	6·75	N.L.	2·9	4·6	5·2	6·4
	3·5	3·1	3·4	4·75	A.L.	3·1	3·9	3·4	4·4
	3·6	5·25	5·75		S.L.	3·6	5·0	6·25	
RM	3·25	4·25	5·1	6·5	N.L.	3·1	4·3	5·0	6·4
	3·0	4·0	5·4	5·4	A.L.	2·7	4·0	5·9	6·4
	3·1	3·25	4·5		S.L.	3·2	3·6	4·3	
EL	3·4	4·6	5·25	4·9	N.L.	3·2	4·9	4·25	5·0
	3·6	4·25	5·0	4·6	A.L.	3·5	4·7	5·6	5·25
	2·75	4·4	4·9		S.L.	2·8	4·4	5·25	
CB	3·9	4·1	4·6	4·4	N.L.	3·4	5·1	4·9	4·5
	3·25	4·75	4·4	4·25	A.L.	3·3	4·6	4·75	4·8
	3·0	4·6	5·0		S.L.	3·8	4·8	4·4	
BS	3·25	3·6	4·4	4·5	N.L.	3·9	4·9	5·75	6·75
	1·25	2·9	3·6	3·6	A.L.	0·9	2·7	5·25	5·25
	2·1	3·4	4·1		S.L.	2·5	5·25	5·75	
$\bar{X}$	3·2	4·1	4·6	4·8	N.L.	3·0	4·5	4·7	5·5
	2·8	3·7	4·1	4·3	A.L.	2·6	3·8	4·8	4·0
	2·8	4·1	4·6		S.L.	3·1	4·3	5·1	

Tables I and II describe performance on the Main Task scored by each of these procedures. The scores are given in terms of the average number of items correctly reported as a function of the main conditions of the experiment.

The assumption that if the subject knows an item he knows its position fits this data well. A corollary of this is that the difference between scoring procedures does not interact with the main effects in the experiment. An approximate test of this assumption may be made in the following way. Let us assume that the data scored for position correct (Table I) represents all the information available to the subject. (In fact a subject randomly guessing from the alternatives would average less than half an item in the correct position.) Let us further assume that having transmitted this information the subject fills the remaining cells with items drawn randomly from the remaining alternatives. Finally, let us assume that the subject is oblivious of the existence of an acoustically similar subset of possibilities or unable to identify occurrences of that set. The number of items regardless of position which such guessing might add to performance was computed for AS and AN stimuli under No Load and Auditory Load conditions at exposure durations of 100, 400 and 1250 ms. These data were compared to the actual obtained

TABLE II

Average number of consonants reported correctly irrespective of position

Subject	Acoustically similar				Exposure (ms)	Acoustically neutral			
	100	400	1000	1250		100	400	1000	1250
S.S.	4·9	5·5	5·1	5·6	No Load	3·8	4·3	6·1	5·75
	4·5	5·4	5·5	5·4	Auditory Load	3·9	4·5	5·4	5·6
	5·4	5·4	5·25		Spatial Load	3·9	5·1	5·7	
AB	5·0	6·4	6·75	7·6	N.L.	5·0	6·4	6·7	7·3
	4·5	5·0	5·5	7·0	A.L.	5·0	5·4	5·1	6·0
	5·1	5·9	6·5		S.L.	5·4	6·25	6·6	
RM	4·5	5·25	6·0	7·0	N.L.	4·6	5·6	6·2	6·9
	4·25	5·1	6·4	6·6	A.L.	3·6	5·4	6·5	7·0
	4·0	4·6	5·6		S.L.	4·6	5·25	5·1	
EL	5·6	6·0	6·75	6·75	N.L.	5·1	6·1	6·1	6·5
	5·25	5·25	6·6	6·5	A.L.	5·3	6·2	6·9	6·9
	5·0	5·4	6·25		S.L.	4·4	5·8	6·3	
CB	5·6	6·0	6·4	6·4	N.L.	5·2	5·9	6·2	5·8
	4·0	6·0	5·4	6·4	A.L.	5·1	6·25	6·7	6·8
	4·9	6·0	6·0		S.L.	5·6	6·25	6·6	
BS	4·75	5·1	6·1	5·9	N.L.	5·1	6·1	7·1	6·4
	2·6	3·4	4·9	5·75	A.L.	3·7	4·6	5·9	6·1
	3·75	4·6	5·1		S.L.	4·8	6·3	6·2	
X̄	5·1	5·7	6·2	6·9	N.L.	4·8	5·7	6·4	6·5
	4·2	5·0	5·7	6·3	A.L.	4·4	5·4	6·1	6·4
	4·7	5·3	5·8		S.L.	4·8	5·8	6·1	

improvement. The predictions were generally quite accurate. Averaged over durations they ranged from an under-estimation of 0·3 items in the NL/AS condition to an over-estimation of 0·2 items in the AL/AS condition. Invariably the predictions changed from over to under-estimation of the improvement as exposure duration increased. Since by chance the proportion of guesses which are correct with respect to position will increase with performance level due to the decreasing number of "free" cells the amount of improvement should be over-estimated by our approximate model. It seems to be the case therefore that part of the improvement with increased exposure duration consists of items retrieved without positional information. This finding echoes that of Henderson (1972), albeit in a very different experimental paradigm.

We may now concede that the assumptions about guessing are not strictly accurate. It is evident from an examination of the raw data that the subject does not guess optimally. Some repetition of items in the response have been noted. In addition there were occasional intrusions of vowels in the responses. Further-more, there was occasional evidence of correct items in the wrong positions due to simple transpositions. These were most commonly mirror-image transpositions but some transpositions between rows also occurred.

Virtually all the features of this analysis conflict with the notion that recall in this task is from immediate verbal memory. The accuracy of position performance is not to be accounted for by the conversion of spatial position to serial order by an implicit speech mechanism. The idiosyncracies of performance from trial to trial are incompatible with the notion of a systematic scan of the matrix. In addition when performance is plotted as a function of spatial position no regular spatial vector can be extracted which yields the ubiquitous serial order function of auditory STM. Again, the effect of acoustic confusability on auditory STM is generally to impair order performance while leaving item performance unimpaired or even improved (Baddeley, 1968). No such systematic increase in the number of items correct but in the wrong position was to be found for AS stimuli.

Since apart from a slight interaction with exposure duration the scoring pro-cedures did not interact with the experimental conditions the remaining analyses will be in terms of items scored for correct position.

There was no evidence of an effect of Spatial Load. When performance is averaged over all conditions the results for the SL condition are 4·00 items and for NL 4·02 items. Neither acoustic confusability nor exposure duration interact with Spatial Load. Table III shows that subjects average 4·8 locations correct out of a possible 6 in the Spatial Task. Despite the conservative scoring procedure performance is high. Subject AB was able to average 5·5 locations correct in the Spatial Task, while being the best performer in the Main Task. The conclusion is inescapable that the processing demands of the two tasks are independent despite their common concern with spatial information.

The Spatial Task was originally chosen because of its apparent lack of verbal codability and its use of non-verbal reproduction at recall. Such tasks are difficult to find and the user has to fall back on face validity in interpreting the task's processing demands. That the elements in the task are codable into multi-location "chunks" is clearly evident from the high-level of performance, from the

TABLE III

Mean number of items correctly located in the Auditory and Spatial Tasks and total number of items correctly recalled in the Auditory Load Condition

Subjects	Auditory Task					Exposure (ms)	Spatial Task			
	100	400	1000	1250	X̄		100	400	1000	X̄
SS	3·0	2·8	3·0	3·2	3·0		4·1	5·0	5·4	4·6
AB	5·3	6·2	6·3	8·0	6·5		5·7	5·6	5·2	5·5
RM	4·0	4·7	5·3	6·7	5·2		4·7	4·8	5·2	4·9
EL	6·0	6·0	6·3	7·0	6·6		5·0	5·3	4·9	5·1
CB	6·2	6·1	6·1	5·8	6·1		5·0	4·6	4·3	4·6
BS	5·0	5·3	4·0	4·5	4·7		3·6	3·7	4·8	4·0
X̄	4·9	5·2	5·2	5·9	5·3		4·7	4·8	5·0	4·8

Main Task

(Auditory Load)

	100	400	1000	1250
X̄	2·6	3·8	4·8	5·0
Total digits + consonants	7·5	9·0	10·0	10·9

subjects' introspective reports and from the systematically transpositional nature of errors.

Turning to the Auditory Load condition we find a different picture. Figure 1 allows the comparison of the AL and NL conditions for both sets of stimulus material There is a small effect of carrying the Auditory Load. The performance decrement is generally about half an item and does not vary systematically with exposure duration or acoustic confusability. From Table III it is evident that the loss of half an item in the visually presented task is accompanied by the successful recall of an average of 5·3 random digits in their correct serial order. The total number of alphanumeric items successfully transmitted rises from an average of 7·5 at 100 ms. exposures to 10·9 at 1250 ms.

From Table I and Figure 1 it is also clear that acoustic confusability generally has a slight deleterious effect on performance. This effect does not vary with load condition but it does interact with exposure duration, the effect increasing with exposure duration or with performance level.

The effect of acoustic confusability is interesting for several reasons. It is smaller than would be expected in Auditory STM (Sperling and Speelman, 1970; Baddeley, 1968), and it is not attributable to a loss of order information since the effect is of the same magnitude when scored for items correct regardless of position (Table II).

A deleterious effect, though small, is also rather surprising when we consider that the AS stimuli occurred with a frequency of one in three and were invariably the same, readily classified letters varying only in order. In the course of his labours the patient subject encountered that set of letters about 100 times, yet no
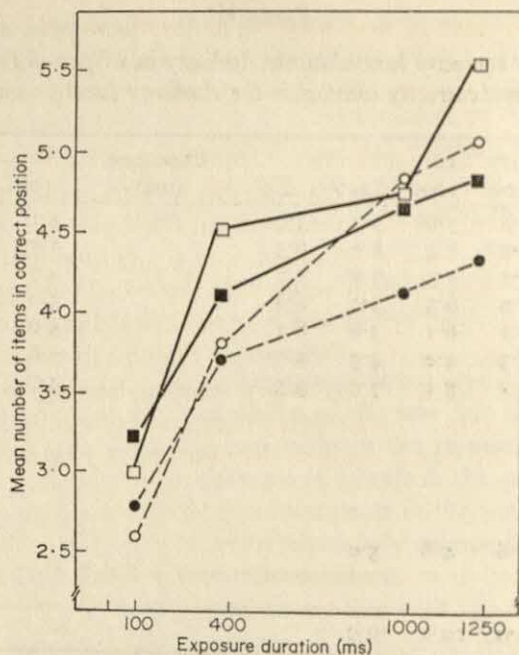
FIGURE 1. Performance as a function of exposure duration for Auditory Load (circles) and No Load (squares) conditions. Open and filled symbols denote Acoustically Neutral and Confusable stimuli, respectively.

subject could give any account of a recurrent, acoustically distinguished set of items. Similar failures to obtain an effect of repetition have been reported by Turvey (1967) using briefer exposures.

The constancy of the AS effect over load conditions suggests in a circumstantial way that any role of naming in the encoding of items does not vary greatly between the load conditions. The interaction whereby AS produced deficit increases with exposure duration is susceptible to a variety of interpretations. It might be claimed, for example, that a naming mechanism increasingly supplements a non-speech store at longer exposure durations (Scarborough, 1971) and this naming mechanism is sensitive to AS. This seems unlikely due not only to the incompatibilities between the AS effect and a verbal STM code already mentioned but also because the AS effect is unchanged in the Auditory Load condition when naming should be impossible.

Concerning the effect of the Auditory Load on the tachistoscopic task the central question is whether an implicit speech mechanism is wholly excluded as a mediating code. The first line of evidence beyond the smallness of the AL effect itself is the size of the combined span (Table III). The average of 10·9 items in their correct positions at 1250 ms exposures clearly exceeds the magic number seven plus or minus two (Miller, 1956) which characterizes the span of immediate verbal memory. Furthermore, it is not the case that a dual series of digits and consonants by its categorical diversity leads to a greatly enlarged combined span. Sanders and Schroots (1969) studied the recall of 6 digits followed by 6 consonants, the converse

list and 12 consonant and 12 digit lists.   The lists were presented auditorily with a slight break halfway, and were scored for items in the correct serial position. Span increased only from an average of 5·0 items in the homogeneous lists to 6·25 in the split lists.

Furthermore, the total span estimated from the block averages probably underestimates maximum performance considerably.   This is due to a PI effect which builds up for the digit series (but not the visual consonants).   This effect accumulates rapidly achieving a maximum about trial 5.   On the first trial of each block the subjects achieve about 2 digits more than the overall average, for example.

The subject has some slight control over the hiatus between the auditory digits and his initiation of the visual presentation.   Perhaps in this interval of about 2 s he hastily tucks the digits into *secondary memory*, leaving *primary memory* free to receive the consonants.   This argument can never be refuted but it encounters difficulties in Table IV which displays the outcome of a serial position analysis of

TABLE IV

Per cent correct at each serial position in the Auditory Task

Subjects	1	2	3	4	5	6	7	8	9
SS	20	16	14	24	15	22	40	51	97
AB	85	60	52	62	63	75	75	63	96
RM	74	58	70	51	41	53	51	44	81
BS	57	51	54	50	66	61	23	26	91
EL	70	61	63	61	56	61	83	86	93
CB	65	64	64	63	64	68	56	65	80
X̄	62	52	54	52	51	57	55	56	91

the Auditory Task.   The feature of interest here is the substantial recency effect. For every subject the final position is recalled best and the average magnitude of this effect is about a 50% improvement on performance at the other positions. The only account of primary memory which can be bent to these data would appear to be that of Crowder and Morton (1969).   Such an interpretation would commit us to the assumption that P.A.S. mediates the recency effect for auditory digits and is unaffected by the acquisition of tachistoscopic information by an implicit speech code and written report from that code.

With the present procedure no auditory control was considered satisfactory since it is impossible to simulate the interpolation of the tachistoscopic task while simply avoiding the actual memory load.   Data gathered in the fourth session suggest that auditory performance is about one item better when the subject only goes through the motions of the tachistoscopic task without receiving and reporting a stimulus. It is possible, therefore, that the subject trades a little auditory performance for superior visual performance.   However, it is equally likely that the *report* of nine consonants interferes with auditory digit performance rather than the demands of storage.

Tachistoscopic report as a function of exposure duration improves in the negatively accelerated manner typically found.   After 400 ms performance settles to a

slow 1 item/s improvement.   It is difficult to see why an implicit speech mechanism with an operating rate of 4–6 syllables/s should yield such a slow improvement, yet a similar slope is reported by Sperling (1960) and Scarborough (1971).

One possibility is that the preservation of items in the visual code demands attention or general processing capacity which detracts from the rate of naming. Another possibility is that visual and name codes acquire information independently with consequent overlap.   This assumption would allow us to double our estimate of the rate of implicit speech but the estimate would remain surprisingly low. However, conversely, it may be that further items are acquired by the visual code more slowly than the first four.   This rate of acquisition might become negligible when primary verbal memory is as extended as in Scarborough's study, due to the amount of general processing capacity absorbed by the visual code.   An unconditional limit of about four items on the visual code is difficult to accept not only in view of the lack of effect of the Spatial Load but because of the difficulty of giving a functional rationale for such a limit.   But then it is perhaps difficult to give a naturalistic rationale for a single glimpse at random consonants.

Conclusion

The notion of an immediate verbal memory with structurally limited capacity to which items must be transferred from iconic memory does not seem tenable. If this is so then performance in tachistoscopic report tasks cannot be wholly accounted for by reference to the properties of this auditory–verbal store.   There must exist a manner of storing visual information which preserves certain spatial information and which is sufficiently independent from the naming–auditory storage loop to provide minimal interference with auditory–verbal memory.

We may call this storage system a *visual code* in acknowledgement of its preservation of spatial information and its ability to be searched spatially.   To be visual in this sense is not necessarily to be *visible* as an image, nor is the modality of the original stimulus necessarily vision.   Such dependencies as exist between the visual and verbal codes may be accounted for by assuming, as Posner (1969) does, that the codes share central processing capacity.   This is not at all equivalent to competing for a fixed memory span.

The concept of relatively independent visual and verbal codes is not in itself sufficient to account for Sanders and Schroots (1969) continuous effect of categorical diversity nor does it account for the lack of effect of Spatial Load on tachistoscopic report in the present experiment.   It is clearly unsatisfactory to attempt an explanation of these results by proliferating hypothetical codes like some array of discrete receptacles.   (Visual "stores" are even more prone to hypostasis than "codes".)   A more satisfactory resolution is to allow that any dimensional structure in stimulus lists may be used to index the material for retrieval.   Visual and verbal codes may be viewed as two areas in coding space which are relatively distinct or discriminable in terms of their indexing.

References

ATKINSON, R. C. and SCHIFFRIN, R. M. (1968). Human memory: a proposed system and its control processes. In SPENCE, K. W. and SPENCE, J. T. (Eds), *The Psychology of Learning and Motivation. Vol.* 2, pp. 89–195. New York: Academic Press.

BADDELEY, A. D. (1968). How does acoustic similarity influence short term memory? *Quarterly Journal of Experimental Psychology*, **20**, 249–64.

COLTHEART, M. and GEFFEN, G. (1970). Grammar and memory. I. Phonological similarity and proactive interference. *Cognitive Psychology*, **1**, 215–24.

CROWDER, R. G. and MORTON, J. (1969). Precategorical acoustic storage (P.A.S.). *Perception and Psychophysics*, **5**, 365–73.

GLANZER, M. and CLARK, W. H. (1963). Accuracy of perceptual recall: an analysis of organization. *Journal of Verbal Learning and Verbal Behavior*, **1**, 289–99.

HENDERSON, L. (1972). Visual and verbal codes: spatial information survives the icon. *Quarterly Journal of Experimental Psychology*, **24**, 439–47.

HENDERSON, L., COLES, S., MANHEIM, M., MUIRHEAD, J. and PSUTKA, P. (1971). Orientation specific masking of letter features. *Nature*, **233**, 498–9.

MILLER, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, **63**, 81–97.

NEISSER, U. (1967). *Cognitive Psychology*. New York: Appleton-Century-Crofts.

POSNER, M. I. (1969). Abstraction and the process of recognition. In BOWER, G. H. and SPENCE, J. T. (Eds). *Psychology of Learning and Motivation. Vol.* 3. New York: Academic Press.

POSNER, M. I., BOIES, S. J., EICHELMAN, W. H. and TAYLOR, R. L. (1969). Retention of visual and name codes of single letters. *Journal of Experimental Psychology Monograph*, **79**, 1–17.

SANDERS, A. F. (1968). Short term memory for spatial positions. *Psychologie*, **23**, 1–15.

SANDERS, A. F. and SCHROOTS, J. J. F. (1969). Cognitive categories and memory span, III. Effects of similarity on recall. *Quarterly Journal of Experimental Psychology*, **21**, 21–8.

SCARBOROUGH, D. L. (1971). Memory for brief visual displays: the role of implicit speech. Paper presented to the Eastern Psychological Association, New York.

SPERLING, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, **74**, 1–29.

SPERLING, G. (1963). A model for visual memory tasks. *Human Factors*, **5**, 19–31.

SPERLING, G. (1967). Successive approximations to a model for short term memory. *Acta Psychologica*, **27**, 285–92.

SPERLING, G. and SPEELMAN, R. G. (1970). Acoustic similarity and auditory short-term memory: experiments and a model. In NORMAN, D. A. (Ed.), *Models of Human Memory*. New York: Academic Press.

TURVEY, M. T. (1967). Repetition and the preperceptual information store. *Journal of Experimental Psychology*, **74**, 289–93.

PERSONALITY IN MONKEYS: FACTOR ANALYSES OF RHESUS SOCIAL BEHAVIOUR

A. S. CHAMOVE,† H. J. EYSENCK,‡ AND H. F. HARLOW

Regional Primate Research Center, University of Wisconsin

Three factor analyses were performed on social interaction data from 168 juvenile macaques. Animals were tested in stable quadrad peer groups; in newly-formed dyads with infant, juvenile, and adult stimulus monkeys; and in similar triads with the stimulus animal plus a familiar cage-mate. Factors emerged, most strongly in the most stable condition, which were interpreted as affiliative, hostile and fearful. These factors were almost entirely independent and resembled the extraversion, psychoticism, and emotionality factors frequently found in humans.

Introduction

Factor analyses of human social behaviour have been undertaken primarily in two areas. The first is upon humans in a solo setting, usually having the subject fill in a questionnaire. The subject is rating his own behaviour in another postulated setting, usually an interacting one. Reliably, factors of introversion-extraversion, emotionality or neuroticism, and psychoticism emerge (reviewed recently by Eysenck and Eysenck, 1968, 1969). Looked at from the point of view of the behaviour patterns on which these factors are based, extraversion is characterized by sociable behaviour, neuroticism by fearful behaviour, and psychoticism by hostile, aggressive behaviour. The second area is upon humans in an interactive setting, usually dyadic. Studies of mother and child (Becker and Krug, 1964; Hatfield, Ferguson and Alpert, 1967), adults (Borgatta, 1964; Schaefer, 1959), children (Borgatta and Sperling, 1963), and infants (Cobb, Grimm, Dawson and Amsterdam, 1967) report factors which can be termed affection or extraversion, emotionality or neuroticism, and assertive reponse to assertiveness. One might expect to find similar patterns of behaviour in the higher subhuman primates. Certainly when primatologists describe behaviour, they often do so using comparable terms.

In subhuman primates two studies have been published, both in the latter category. The first (Locke, Locke, Morgan and Zimmerman, 1964) recorded 10 behaviour categories in 12-year-old rhesus monkeys that had been reared in social isolation (probably not visual or auditory isolation) from soon after birth. The severely restricted repertoire of isolated monkeys, when tested in quadrads, enabled

†Present address: Stirling University Psychology Primate Unit, Stirling, FK9 4LA, Scotland.
‡Present address: Institute of Psychiatry, University of London, London, S.E.5., England.

the experimenters to record only the following behaviours: pass, approach, contact, chase, aggression, passive awareness, avoidance, escape, submission, and apparent unawareness. Two factors emerged: dominance, and submission in response to dominance.

Van Hooff (1971) reported a component and cluster analysis of 53 behaviours recorded in a stable group of 25 chimpanzees. Sixty-nine per cent of the variance was accounted for by components termed affinitive or social positive, play, aggressive, and submissive. In addition, contributing a small but significant amount to the results, were factors termed groom, excitement, and "show" or display.

The following is a report of three factor analyses, of 10 behaviours, performed on data from a stable group situation and on data from less stable dyadic and triadic situations.

Method

Subjects

One hundred and sixty eight *Macaca mulatta*, about 85% males, were separated from their mothers at birth and reared in individual mesh cages. They were given daily peer experience starting at between 15 and 90 days of age. For about half the animals this consisted of daily 2-h pairings in the homecage and bi-weekly 1-h group sessions in a playroom or in a large cage. For the rest social experience entailed daily 40-min group sessions in a playroom. All subjects were assigned to a group composed of four age-mates, and all social experience, both pairings and group sessions, involved these group members. All group social interaction was experienced when all four members were present, so that after a few months the group had formed stable social relationships which were retained even when monkey subjects were paired.

Apparatus

The animals were group tested in that particular playroom or large cage to which they had become accustomed through daily group sessions throughout most of their lives. The playrooms used were Wisconsin Playroom II for 44 animals (see Chamove, Waisman and Harlow, 1970, for a description), Wisconsin Playroom III for 36 animals (described by Chamove, 1966), and a standard rectangular mesh cage measuring $3 \times 1 \cdot 5 \times 1 \cdot 5$ m for the remaining monkeys. Stimulus testing was carried out in a similar sized cage but fitted with a plexiglas front and a plexiglas partition which could divide the cage in half. It allowed the animals to observe one another in the adaptation interval before testing.

Behaviours were recorded using a bank of 10 Standard Electric Timers activated by 10 microswitches (as described by Chamove, Harlow and Mitchell, 1967). The timers recorded the duration of the following behaviours: *social explore*—any investigation, contact or not, of another animal, primarily looking at another monkey (reliability, $r = 0 \cdot 88$); *social play*—playing with another monkey, usually a relaxed, complex, and vigorous behaviour ($r = 0 \cdot 95$); *nonsocial play*—similar behaviour not directed toward another animal ($r = 0 \cdot 90$); *nonsocial fear*—withdrawal from the environment, scored when no social object could be credited with the instigation of the fear response ($r = 0 \cdot 96$); *appropriate withdrawal*—avoidance of an animal that is exhibiting hostile behaviour ($r = 0 \cdot 91$); *inappropriate withdrawal*—withdrawal from an animal exhibiting fear, exploratory, or play behaviour ($r = 0 \cdot 90$); *hostile contact*—biting or grabbing another animal ($r = 0 \cdot 96$); *nonhostile contact*—all other behaviours involving physical contact except clinging ($r = 0 \cdot 95$); *social cling*—clinging to another animal ($r = 0 \cdot 99$); *noncontact hostile*—($r = 0 \cdot 91$).

Reliability is given as product-moment inter-observer coefficients.

Procedure

Group Testing was done on all subjects between 9 and 12 months of age, and all monkeys were tested in their group of four a minimum of twice weekly. Testing consisted of at least thirty 60-min sessions during which each animal was observed for three 5-min periods. One experimenter did all the Group Testing. The data were converted to "per cent of total time tested" scores for purposes of analysis.

Stimulus Testing was performed on 63 of the above subjects soon after Group Testing. At 12 months of age the monkeys were housed in groups of four continuously. At approximately 13 months of age the first Stimulus Test was undertaken. The subject was housed in one half of the Stimulus Test cage, separated from the stimulus animal by the plexiglas partition, for a 23-h period. The transparent barrier was then removed and testing immediately begun; it was continued for 15 min unless one of the animals was seriously injured.

Testing involved placing the subject animal with a stimulus animal of relatively constant characteristics in the following 3 × 2 design. The stimulus monkey was either an infant of approximately 1 month, a juvenile male of approximately 6 months of age and chosen for its playfulness, or an adult male chosen for its docility. The subject was tested with each of these three in the above order in a dyadic social situation. The tests were separated by a period of about 10 days. After the "Alone" Stimulus Test, another three sessions followed. The subject monkey was first paired with an animal from his original four-membered test group. The two animals were quite familiar with one another, having lived together for over a month in a group of four and having interacted with one another from an early age. These two monkeys were then tested *together* with the stimulus animal. Here two experimenters were used. The procedure followed was the same as for the Alone testing except that two animals were adapted for 23 h in half of the test cage, and the test situation was triadic—two subjects and one stimulus monkey. This was called the "Partner" situation.

Analysis of the data was by means of principal components analysis rotated to oblique simple structure (Eysenck and Eysenck, 1969). Interfactor correlations of the promax factors were derived and all factors with eigenvalues less than unity were ignored. Computer limitations forced a division of the Stimulus Test data into two halves, the Alone and the Partner testing situations. As clinging was rarely observed in the Group Test situation, it was deleted from the analysis.

The rationale of the procedure is as follows. The Group Test situation involves a maximum of adaptation to the social test situation. Because of the formation of hierarchies and preferences during their year of social interaction, one would expect a maximum of within-group divergence, the most subordinate subject, e.g. being able to show almost no hostility. Tempered by this very stable structure, personality differences might be expected to emerge.

The Stimulus situation has been shown to demonstrate the effects of early environmental manipulations which the group situation has failed to detect (Mitchell, 1970; Chamove, 1966), but it may be less reflective of habitual modes of social response due to the variability of the various test situations. Dominance position, dyadic alliances, and social response are much more controlled in this stimulus situation. The stimulus infant at 1 month moves about very little and normally elicits very little aggression. The juvenile is very active but always subordinate to the experimental subject, and so aggression is more commonly elicited; play is quite common and fear uncalled for. The docile adult is a test for aggressive-fearful factors. Excessive fear should be rare and aggression rarer, brief submissive gestures being more appropriate with these males. Isolate monkeys, however, suicidally attack all animals (Mitchell, 1970) whereas certain brain damaged monkeys carefully discriminate the objects of their hyper-aggression (Chamove et al., 1970).

The Partner situation is used because it was felt that having a familiar peer present might elicit certain behaviours not elicited in the Alone condition as has been previously found (Chamove, 1966).

Results

The results of the factor analysis of the *Group Test* data are striking. Three clear factors emerge having little intercorrelation: hostile, fearful, and affiliative or sociable (see Table I). These three all show positive intercorrelations. This may be a result of the constrained social situation, for an animal showing play or hostility may induce fear in other animals. Another possible cause of this correlation may be the existence of a kind of monkey which is relatively nonsocial and inactive; he will score low on all behaviours. The factor we have here termed hostility correlates 0·23 with the one we termed fear, 0·34 with affiliation while the affiliation factor correlates 0·21 with fear. These correlations account for only 10% of the variance at most, and we consider them unimportant.

TABLE I

Loading of nine behaviours on three factors resulting from analysis of the quadrad Group Test data and labelled play, fear, and aggression–hostility

	Factors		
Behaviours	I	II	III
Nonsocial play	0·10	0·78	0·02
Social play	−0·07	0·99	0·03
Positive contact	−0·04	0·96	−0·01
Social exploration	0·73	0·39	−0·15
Nonsocial fear	0·99	−0·16	0·04
Inappropriate fear	0·91	0·04	−0·03
Appropriate fear	0·95	0·00	0·14
Noncontact hostility	0·40	−0·10	−0·80
Contact hostility	−0·25	0·04	−0·98

One nonfear variable, noncontact hostility, shows a moderate correlation with the fear factor. This single aberrant score can be explained in terms of the behaviour pattern termed redirected-threats. These are hostile behaviours directed toward a more subordinate animal in an attempt to distract the hostile attentions of a more dominant monkey. The behavioural sequence is initiated by the dominant animal and the redirection is a response to that initiation.

The results of the *Stimulus Test* are less clear, as might be predicted from the unstable nature of the test situation (see Table II). When monkeys are tested in a dyad with the stimulus animal (upper half of Table II), four factors are found. The first loads negatively on infant fear, juvenile fear and hostility, and adult hostility; and positively on infant hostility. The second factor is one of juvenile play versus infant positive contact, probably of an exploratory nature. The third loads on nonsocial fear versus social explore. Both are fairly nonsocial behaviours as they involve little social contact and are seen in nonsocial animals such as long-term isolates. The final factor is social play versus adult hostility.

When monkeys are stimulus tested paired with familiar partners four factors again emerge (lower half of Table II). The first loads positively on play and on cling directed toward the juvenile and adult stimulus animal versus a negative

TABLE II

Loadings on the first four factors resulting from analysis two (upper) and analysis three (lower) of Stimulus Test data when subjects are tested with three types of stimulus monkey

| | | Factors | | | | | | | | | | | |
| | Behaviours | I | | | II | | | III | | | IV | | |
		infant	juvenile	adult	infant	juvenile	adult	infant	juvenile	adult	infant	juvenile	adult
"alone" or dyad condition	Nonsocial play	0·04	0·05	0·27	0·16	0·03	0·15	0·22	0·17	0·15	0·02	0·36	0·14
	Social play	0·18	0·23	0·16	−0·06	−0·51	−0·09	0·56	0·22	0·07	0·23	0·27	0·26
	Positive contact	0·10	−0·24	0·01	0·75	0·48	0·03	0·09	0·00	0·09	−0·13	−0·09	0·21
	Social cling	0·20	−0·36	−0·11	−0·03	−0·26	−0·29	−0·05	0·01	0·04	−0·09	0·05	0·10
	Social exploration	−0·13	0·01	0·18	0·49	0·35	−0·20	−0·08	−0·21	−0·04	−0·33	0·09	−0·48
	Nonsocial fear	0·28	0·04	0·15	0·03	−0·05	0·12	−0·17	0·06	0·04	0·17	0·52	0·45
	Inappropriate fear	−0·55	−0·64	−0·28	0·09	0·12	0·09	0·17	0·21	−0·86	0·15	−0·04	−0·11
	Appropriate fear	0·00	−0·68	−0·37	0·00	0·04	0·07	0·00	0·11	0·08	−0·00	−0·06	−0·34
	Noncontact hostility	−0·02	−0·48	−0·61	−0·01	0·06	−0·08	0·10	−0·23	−0·13	0·17	−0·09	−0·07
	Contact hostility	0·73	−0·33	−0·69	0·06	−0·13	−0·04	0·01	−0·33	−0·19	0·18	0·08	0·12
"partner" or triad condition	Nonsocial play	−0·40	−0·48	0·19	−0·05	0·02	0·31	0·07	0·15	0·48	0·05	0·11	−0·09
	Social play	0·05	0·52	0·24	−0·73	0·16	−0·08	0·37	−0·14	−0·06	0·18	0·44	−0·42
	Positive contact	0·04	0·68	−0·14	0·62	0·11	−0·03	0·04	0·01	−0·13	0·07	0·01	−0·41
	Social cling	0·10	0·50	0·33	−0·06	−0·02	−0·04	0·46	0·21	0·12	0·14	−0·27	−0·08
	Social exploration	0·04	0·44	−0·45	0·70	0·08	0·19	0·01	0·14	−0·01	0·09	−0·25	0·19
	Nonsocial fear	−0·05	−0·05	0·05	0·28	−0·00	−0·22	−0·09	0·15	0·03	−0·38	−0·03	0·12
	Appropriate fear	−0·36	0·05	−0·06	−0·16	−0·16	−0·13	−0·04	−0·42	−0·11	−0·20	0·32	0·15
	Noncontact hostility	−0·10	−0·24	−0·07	−0·65	−0·12	−0·21	−0·35	0·05	−0·65	0·07	0·07	−0·11
	Contact hostility	−0·14	−0·06	0·18	−0·42	−0·15	0·21	−0·08	−0·05	−0·82	0·02	0·06	−0·20

loading on nonsocial play and social fear. The second loads on infant hostile and play versus infant explore and nonsocial fear. The third factor loads on adult and infant hostile, juvenile fear versus infant play and cling and adult nonsocial play. The last factor loads positively on juvenile play and inappropriate withdrawal as contrasted with negative loading on adult play, juvenile explore and cling, and infant nonsocial fear. Table III lists the intercorrelations between each of the four factors from both of the Stimulus Test factor analyses. Except for factors I and II, the correlations between pairs of factors is remarkably low suggesting a high degree of independence.

TABLE III

Interfactor correlations of promax factors from Stimulus Testing

Analysis two; dyads			
Factors	II	III	IV
I	0·33	−0·27	0·07
II		−0·18	−0·09
III			0·06
Analysis three; triads			
Factors	II	III	IV
I	0·61	−0·12	−0·08
II		−0·17	−0·04
III			−0·03

The complexity of the Stimulus Test data is what one might expect if for no other reason than that different behaviours are elicited in the six situations. For example, in the Partner or triad situation the subjects exhibit less juvenile-directed hostility, less inappropriate fear toward the adult, more adult hostility, and less nonsocial fear, substituting partner-directed clinging in many cases for nonsocial fear. Factors III and IV, found in the Alone condition, are represented in factor I of the Partnered situation; factors II and III of the Partner test are found in IV and I respectively of the Alone condition.

The first factor of the Alone condition and second factor of the Partner test are clearly a social fear and hostile category. The second factor of the dyadic test appears comparable to an affectional category, showing play toward the juvenile and restrained positive and exploratory behaviour toward the infant, the first infant these monkeys had ever seen. The first factor of the triad or Partner test, like factors III and IV on the dyad test, seem also to reflect an affectional character, perhaps more comparable to the meaning of extraversion. We see here an animal exhibiting play toward the juvenile and other monkeys, showing some avoidance of other animals, considerable nonsocial play, and some hostility toward the playful juvenile stimulus monkey. The last triad factor dichotomizes between adult play and juvenile play.

Discussion

The factor analysis of the Group Test has resulted in three clear behaviour factors: fearful, hostile, and affectionate. These patterns of behaviour are not dissimilar to those which gave rise to the three major factors in research on human personality; neuroticism-stability, extraversion–introversion, and psychoticism. It would be premature to seek to prove the identity of the factors in these different species; no acceptable method exists at the moment for any such proof. What would be required to make the identification reasonable would be the incorporation of these factors in a nomological network, rather in the fashion adopted by Eysenck and Broadhurst (1964) with reference to emotionality in rats. Thus it should not prove impossible to test whether affectionate (extraverted) monkeys have lower cortical arousal patterns than do non-affectionate ones; similarly, it should be possible to test for differences in conditioning between the two groups. A programme for testing the suitability of the "neuroticism" tag for the fearful animals could with advantage follow that adopted for emotionality in rats. Psychoticism might be the most difficult factor to investigate, for the simple reason that least is known about it in the human population. However, even here such behaviours as lack of co-operation or inappropriate aggression should be susceptible to observation and quantification.

It might at first seem surprising that similar factors emerge from two different species, but there are good reasons for expecting such agreement. The first reason is linked with the simple fact that monkeys and men (and rats also) have similar anatomico-physiological structures to subserve emotional/fearful behaviour, i.e. an autonomic system and a visceral brain, and arousal behaviour, i.e. a cortex linked with an ascending reticular formation; one would expect individual differences in behaviour to be linked with differences in the functioning of both these systems, and these behavioural differences would be expected to be the more similar to those observed in humans, the closer the species under investigation was to *Homo sapiens*. As regards psychoticism, it is much more difficult at this moment to suggest a biological locus for this trait, but a close relation has been observed with masculinity. The possibility of a link with some hormonal secretion related to the sex glands would be supported by recent investigations of rhesus monkeys (Rose, Holady and Bernstein, 1971).

The second reason is that we are here concerned with social behaviour, i.e. how one animal behaves towards another, and the major possibilities of such behaviour seem to be limited to the three patterns we noted; an animal can be friendly-sociable-affectionate, it can be hostile-aggressive-cruel, or it can be fearful-emotional-withdrawing. Most if not all social behaviour can be grouped under these three main headings, and it is perhaps not too surprising that these patterns should emerge as factors in our investigation. Nor is it surprising that the picture is somewhat more complex (and confused) when we turn from the Group Test data to those collected in the Stimulus Test. This would be expected if for no other reason than that different behaviours are inevitably elicited in the six situations by the different "stimulus" patterns.

This observational study demonstrates marked individual differences between monkeys in their social behaviour. These differences are apparently highly

reliable, and characteristic of the animals concerned, and may thus be regarded as aspects of the "personality". Yet experimental work with monkeys, and other animals as well, seldom pays attention to their "personality". We would argue that this omission is a serious one, just as experimentation in human subjects can be very misleading if it leaves out of account personality factors like extraversion–introversion (Eysenck, 1967). In rats a whole host of experimental studies can be shown to be influenced profoundly by differences in emotionality, and quite different theoretical conclusions can be drawn from identical experiments depending on the strain of animals used (Eysenck and Broadhurst, 1964; Eysenck, 1967). The usual process of "averaging" serves simply to relegate such "personality" factors to the error term, which thus becomes unduly swollen and often far exceeds in importance the main effects looked for. Personality factors usually emerge as "interaction" factors, and may serve materially to reduce the size of the error term. It is in this function that we see the main importance of our findings; work with monkeys too should bear in mind the importance of individual differences and use scores on these factors to keep the error term as small as possible. In addition there is of course the possibility that continued work along these lines will increase our knowledge about "personality" factors in animals, and humans as well. In view of the lack of knowledge in this field, such increases would be more than welcome.

References

BECKER, W. C. and KRUG, R. S. (1964). A circumplex model for social behavior in children. *Child Development*, **35**, 371–96.

BORGATTA, E. F. (1964). The structure of personality characteristics. *Behavioural Science*, **9**, 8–17.

BORGATTA, E. F. and SPERLING, B. C. (1963). A note on the stability of peer judgements in independent situations. *Journal of Psychological Studies*, **14**, 45–8.

CHAMOVE, A. S. (1966). The effects of varying infant peer experience on social behavior in the rhesus monkey. Unpublished M.A. thesis, University of Wisconsin.

CHAMOVE, A. S., HARLOW, H. F. and MITCHELL, G. A. (1967). Sex differences in the infant-directed behavior of preadolescent rhesus monkeys. *Child Development*, **38**, 329–35.

CHAMOVE, A. S., WAISMAN, H. A. and HARLOW, H. F. (1970). Abnormal social behavior in phenylketonuric monkeys. *Journal of Abnormal Psychology*, **76**, 62–8.

COBB, K., GRIMM, E. R., DAWSON, B. and AMSTERDAM, B. (1967). Reliability of global observations of newborn infants. *Journal of Genetic Psychology*, **110**, 253–67.

EYSENCK, H. J. (1967). *The Biological Basis of Personality*. Springfield: C. C. Thomas.

EYSENCK, H. J. and BROADHURST, P. L. (1963). Experiments with animals. In EYSENCK, H. J. (Ed.), *Experiments in Motivation*. Oxford: Pergamon.

EYSENCK, H. J. and EYSENCK, S. B. G. (1969). *The Structure and Measurement of Personality*. London: Routledge & Kegan Paul.

EYSENCK, S. B. G. and EYSENCK, H. J. (1968).   The measurement of psychoticism: a study
    of factor stability and reliability.   *British Journal of Social and Clinical Psychology*,
    7, 286–94.
HATFIELD, J. S., FERGUSON, L. R. and ALPERT, R. (1967).   Mother-child interaction and the
    socialization process.   *Child Development*, **38**, 365–414.
LOCKE, K. D., LOCKE, E. A., MORGAN, G. A. and ZIMMERMAN, R. R. (1964).   Dimensions
    of social interactions among infant rhesus monkeys.   *Psychological Reports*, **15**,
    339–49.
MITCHELL, G. A. (1970).   Abnormal behavior in primates.   In ROSENBLUM, L. (Ed.),
    Primate Behavior: Developments in Field and Laboratory Research.   London: Academic
    Press.
ROSE, R. M., HOLADAY, J. W. and BERNSTEIN, I. S. (1971).   Plasma testosterone, dominance
    rank and aggressive behaviour in male rhesus monkeys.   *Nature*, **231**, 366–8.
SCHAEFER, E. S. A. (1959).   A circumplex model for maternal behavior.   *Journal of
    Abnormal and Social Psychology*, **59**, 226–35.
VAN HOOFF, J. A. R. A. M. (1971).   *Aspects of the Social Behaviour and Communication in
    Human and Higher Non-Human Primates*.   Rotterdam: Bronder.

ACOUSTIC SIMILARITY AND INTERFERENCE IN SHORT-TERM RECOGNITION MEMORY: SOME RESULTS OF A DECISION-THEORY ANALYSIS†

J. D. INGLEBY

MRC Unit on Environmental Factors in Mental and Physical Illness, London

The following three studies of single-probe recognition memory set out to show the effect on the signal-detectability measures of d' and β (Tanner and Swets, 1954) of variations in the acoustic similarity of interfering material, which may either precede or follow the item to be remembered (proactive or retroactive interference —PI or RI). The first experiment studies a situation employed by Wickelgren (1966*a*), who reported that acoustically similar RI substantially reduced d'. It is shown that this effect could have been due to biases in Wickelgren's original designs, and that when a bias-free design is used, the fall in d' is only of borderline significance.

To investigate this problem further, a design was evolved in which two items were presented for memorizing, which varied in acoustic similarity to each other, and (after a distracting task) a probe was presented with one of three questions: Was this the first item of the pair? Was it the second? or, Did it occur in either position? In the first case, recognition-memory with RI of varying acoustic similarity was being studied, and as in the first experiment, it was found that similarity slightly reduced d'. With the second question, PI effects were being studied, and here negligible differences were found. With the third type of question, a "location-free" test, no effects of similarity were found. The last result rules out Posner's (1967) "acid-bath" explanation of similarity effects in interference: an explanation in terms of "differentiation" (or "filtering") was also invalidated by the results of a third experiment, in which the same effects were found even though similarity varied only between stimulus items and interference, and not between these and the probe. Wickelgren's (1966*b*) associative model appears to have least difficulty in accommodating these results, though even this needs certain *ad hoc* assumptions to be able to do so.

Introduction

The crucial predictions of most theories of memory concern the relationships between certain experimental variables and accuracy of recall. Accuracy of recall, however, is hardly acceptable as a unitary dependent variable since in many tasks false–positive reports contribute to this measure as well as omissions. It is therefore desirable to use some basic model which will enable a single measure of retention to be derived from these two sources of data. Several such models

have been proposed. The one adopted for the situations studied here is based on signal-detection theory (Tanner and Swets, 1954): it is assumed that items have a "trace strength", subject to random Gaussian fluctuations, which is incremented by a certain amount when presentation of an item occurs; the difference between the trace strength of presented and non-presented items, d', provides a measure of retention independent of variations in "response bias", the latter being conceived of either as changes in the absolute value of trace strengths, or in the position of the "criterion" applied to them to determine the response made. The evidence in favour of applying this model to memory tasks has been summarized by Broadbent (1971), though it should be noted that it is not totally distinguishable empirically from some rival models.

In this paper, the d' measure is used to clarify the effects of acoustic similarity in interfering material used in short-term memory tasks, and to attempt to discriminate between three current explanations of these effects. We start by reconsidering a pioneer experiment of this type performed by Wickelgren (1966a).

Experiment I

This experiment replicates one reported by Wickelgren (1966a), with certain modifications. The method he used was to present a single letter to be remembered, followed immediately by transcription of 10 letters and then by a single probe letter, which was either the same as that presented ("old") or different ("new"), but always from the same acoustic class. The interfering material varied in acoustic similarity to the probe and the test letter. Wickelgren employed two versions of this design: in one, the presented letter was always excluded from the intervening material, while in the other it was always included.

If we consider the possible ways of constructing this experiment, it will be seen that there are four main types of design (Table I).

Wickelgren's two versions of the experiment corresponded to designs (a) and (b) in Table I, both of which include inherent biases. In design (a), we see that with similar RI, new probes are included in the intervening material whereas old ones are not. If such inclusion increases the tendency to report an item as "old" (as was found by Ingleby, 1969), the false–positive rate with similar RI will be increased, and d' will accordingly be lowered. Likewise, in design (b), with dissimilar interference old probes are included in the RI while new ones are not, which will raise the obtained d' under dissimilar RI. In a pilot study for the present experiments, both these designs were tested, and similar RI was indeed found to produce lower d' values than dissimilar RI, as Wickelgren's results indicated.

Design (c), in which the variable of inclusion in the interfering material is matched between old and new probes, but not between interference conditions, is still not satisfactory: as well as masking the possible changes in bias produced by similar interference, it may lead to changes in d' if repetition of a presented item in the RI reduces the effective trace increment, as suggested in the work quoted above (Ingleby, 1969). The design which eliminates all these possible artifacts is therefore the fourth one illustrated, in which both probes and presented items

TABLE I

Designs for the RI experiment (see text), showing the cases in which the presented item and the probe are included in the interference, and indicating the source of each possible artifact(*)

(a)

	Presented item		
Similar interference		Dissimilar interference	
Pres. not included		Pres. not included	
Probe OLD	Probe NEW	Probe OLD	Probe NEW
Not incl.*	included*	not incl.	not incl.

(b)

	Presented item		
Similar interference		Dissimilar interference	
Pres. included		Pres. included	
Probe OLD	Probe NEW	Probe OLD	Probe NEW
included	included	included*	not incl.*

(c)

	Presented item		
Similar interference		Dissimilar interference	
Pres. included*		Pres. not included*	
Probe OLD	Probe NEW	Probe OLD	Probe NEW
included	included	not incl.	not incl.

(d)

	Presented item		
Similar interference		Dissimilar interference	
Pres. not included		Pres. not included	
Probe OLD	Probe NEW	Probe OLD	Probe NEW
not incl.	not incl.	not incl.	not incl.

are always excluded from the interference. The adoption of this design brings with it new difficulties. If, to ensure forgetting, it is necessary to present a long string of interfering items, it may become obvious to the subject which letters have been excluded, since the two acoustic classes contain 6–8 items each. With similar interference, one of the missing letters would reveal the identity of the probe. Therefore, to start with, the subjects were made to remember two letters at the outset instead of just one: the additional item, which was merely "ballast", was always dissimilar to the other, and could occur before or after it. This reduced the number of interfering letters necessary to produce a reasonable level of forgetting. However, since the number was still high (*c.* 20 letters), and could still transmit information about the identity of the probe when the latter belonged to the same acoustic class, it was decided in the original version of this experiment (11 subjects) to alternate the interfering letters with digits: this experiment, which is not reported here, yielded no difference between similar and dissimilar interference, so it was decided to use letters alone and take the risk of "information leakage" from the latter to the probe. In the event, the risk was justified, since the direction of the obtained differences was opposite to that which would be expected if subjects were obtaining clues to the correct answer when similar interference was given.

Method

Presentation

Items were presented visually on a "Bina-View" display in the following order, at a rate of about 0·4 s per item:

Letter pair—Shadowing—Probe letter.

A blank frame occurred between the shadowing and the probe letter, but in order to maximize the interfering effect of the shadowing there was no blank between the initial pair and the latter. Subjects initiated each trial themselves, and called out all letters except the probe; they were instructed to memorize only the initial pair.

Stimulus material

The pair of letters to be remembered contained one member from each of the acoustic classes FLMNSX and BCDEGPTV. The order in which these two classes occurred, and the selection of letters from each, were randomized. Meaningful or familiar letter pairs were avoided, using a list chosen by three typical subjects.

Shadowing

The subjects called out a string of letters, which varied in length between 15 and 25 items, depending on the subject's performance level. The letters were all from one acoustic class: they did not include the letter presented in the pair to be remembered or, if an incorrect ("new") probe was to be presented, the probe letter. If the probe was to be correct ("old"), or did not belong to the same acoustic class as the shadowing, another arbitrarily chosen letter was excluded from the shadowing. This procedure was necessary in order that two letters should always be removed from the shadowing, making it equally redundant, and therefore equally distracting, in each experimental condition. For the same reason, the same letter never occurred twice in succession in the shadowing.

Probe

The probe was a single letter chosen from the two classes given above, which was either acoustically similar or dissimilar to the letters for shadowing, and either "old" (presented in the stimulus pair) or "new". These four conditions occurred equally often and at random. Subjects had to indicate by a tick on a four-point rating scale whether they thought the probe was new or old; the divisions on the rating scale were labelled Certain New: Doubtful New: Doubtful Old: Certain Old.

Schedule

Lists of 48 trials were prepared by computer on punched tape for the Bina-View apparatus, with several different randomizations of order and different amounts of shadowing. These lists took about 25 min to complete, marking of answer-sheets taking place every 12 trials. The subjects were nine naval ratings aged 17–21: they each completed four such lists, after one or two practice sessions during which they were encouraged to use each rating category to an equal extent.

Analysis

Detection rates, false-positive rates and values of β were calculated at the central criterion position (the decision between "old" and "new"), while all four confidence ratings were used in the construction of an ROC for each subject. The value of d' was taken from the normalized ROC's thus obtained, in the manner proposed by Clarke, Birdsall and Tanner (1959) and now widely accepted as appropriate; their measure, which they termed de, is two times the absolute value of either co-ordinate of the point of intersection between the fitted ROC and the negative diagonal. (A least-squares method of estimation was used to fit normalized ROC's to the obtained points.) Table II shows the mean results.

<div align="center">

TABLE II

Results of Experiment I

</div>

	Dissimilar interference	Similar interference	P*
Detections	84·9%	75·0%	0·01
False–positives	12·8%	9·7%	—
Log β	0·24	0·69	—
d'	2·22	1·91	—

*P values from Wilcoxon matched-pairs signed-ranks tests, two-tailed. $N = 9$.

Results

The only significant result in Table II is the drop in correct detections produced by similar interference. Unfortunately, it is not possible to decide whether d' or β (or both) are responsible for this: neither of the changes in these measures are significant, though at least one of them must be real.

This ambiguous result seemed to call for further experimentation. Rather than using the same paradigm, however, a design was evoked which enabled the experimenter to compare the effects of acoustic similarity in RI, PI, and a "location-free" test simultaneously. Two items were presented for memorizing, which varied in acoustic similarity to each other; after an unrelated distracting task, a probe was presented with one of three questions: was this the first item of the pair? was it the second? or, did it occur in either position?

Experiment II

Method

Presentation

Each trial consisted of the following string of items, presented visually:

Letter Pair—Digit Shadowing—Probe Letter + Question.

The shadowing followed immediately on the letter pair without a blank frame, so as to reduce the opportunity for rehearsal. Presentation was on the Bina-View display at the rate of about 0·4 s per item. Subjects initiated each trial themselves; they were instructed to call out all items except the probe letter and question, remembering only the initial pair of letters.

Stimulus material

The two letters were chosen from the acoustic classes FLMNSX and BCDEGPTV. The pair was acoustically either similar or dissimilar: the two letters were never identical. These two conditions occurred with equal frequency and in random order. Meaningful or familiar letter pairs were eliminated by the procedure described in Experiment I: this left more dissimilar pairs than similar ones, so the set sizes were equated by further random exclusions.

Shadowing

To discourage rehearsal of the stimulus material, subjects called out a string of random digits (2–9 inclusive). The length of this string varied between 10 and 20 items, according to the level of performance of the subject.

Probe

The probe was a single letter followed by a symbol indicating the type of question to be answered:

    (1)   Was the probe the first letter presented? (RI condition).
    (2)   Was the probe the second letter presented? (PI condition).
    (3)   Was the probe presented at all? (Location-free test).

If a "new" probe was given, it was always a similar letter to the correct one (in conditions RI and PI), but never the same as the other letter in the pair. In the location-free test, the new probe was always similar to one or both of the presented letters. (The reasons for these constraints follow from the issues discussed earlier in the construction of Wickelgren's original experiments.) "Old" and "new" probes occurred with equal frequency and in randomized order, as did the three types of question and the two similarity conditions. This gave 12 different types of trial, which occurred four times each in every list of 48 trials. Answers were given in writing on a four-point rating scale, in the subjects' own time.

Schedule

Several dozen lists were prepared by computer, containing different amounts of shadowing and different randomizations of the 12 types of trial. Each list took about 20 min to complete, and marking took place after each sheet of 24 items. The subjects were 21 naval ratings: they were tested over periods of a fortnight, during which they completed between six and eight full lists, as well as one or more for practice.

Analysis

For each of the three types of tests, detection rates, false-positive rates, and values of $\log \beta$ were obtained at the central position; d' was obtained from the full set of confidence ratings, as described for Experiment I. The differences produced by similarity in the stimulus material, in each of the three test conditions, were then computed: they are given in Table III.

TABLE III
Results of Experiment II (N = 21)

| | RI | | PI | | "Location-free" | |
	Sim	Diss	Sim	Diss	Sim	Diss
% Detections	85·1	87·8	83·4	87·8*	88·4	86·2
%False–positives	16·5	12·6*	14·7	15·1	19·0	21·0
Log β	−0·52	0·02*	−0·43	−0·50	−0·83	−0·94
d'	1·92	2·30†	1·92	1·99	1·98	1·82

Note: "Sim" = similar interference; "Diss" = dissimilar interference; * difference significant at $P < 0.05$ by Wilcoxon matched-pairs signed-ranks test, two-tailed; † difference significant at $P < 0.01$.

Results

Table III shows a fall in both d' and β with similar interference in the RI condition, but no significant changes in these measures in the other conditions. This finding throws light on certain hypotheses that have been put forward to explain interference effects.

One possibility is that when the probe item is similar to the interference list, the subject may have a recollection of its occurrence but be unable to discount the suspicion that it was in the interference rather than the original pair: if the inter-

ference is dissimilar from the probe, he can rule out this source of uncertainty. (A similar explanation was put forward by Brown (1958) to account for the fact that digit interference impairs memory for letters less than alphabetical material, and vice versa; Underwood (1945) also proposed such a concept, under the name "differentiation".) We shall call this the "filtering" hypothesis, following Broadbent (1971): it predicts that retroactive and proactive interference from similar material should both produce a greater fall in d' than from dissimilar material, and that when the subject is not required to say whether an item was in the test material or the interference, but merely whether it occurred at all, d' will be unaffected by the similarity of the interference to the test material.

A second explanation is the "acid bath" hypothesis of Posner (1967): according to this, items in store will decay more rapidly the greater the resemblance between them. This likewise predicts a fall in d' with PI as well as RI, but differs from the "filtering" hypothesis in that it also predicts a fall in d' in the "location-free" test.

A third possibility is the associative theory put forward by Wickelgren (1966b), according to which (in brief) the basic internal units of representation are phonemes or some conceptually similar unit, and serial position information, or other phonemes, are associated to these. (This is in contrast to a "pigeon-holing" or "slotting" type of theory in which different moments in time have different storage locations, and material is associated to these locations.) When two phonemically similar items are presented in close temporal succession, confusion over their respective positions will arise because different positions will be associated with the phoneme-representative they both share. Order-free recall, however, will not suffer, and may actually be improved. Wickelgren (1965) in fact showed that item recall in a list of similar items was better, though order recall worse, than in a list of dissimilar ones. This theory, then, will predict that similarity between item and interference will lower d' whichever occurs first, but predicts no impairment, and possibly an improvement, in a "location-free" test.

The predictions of these theories are summarized in Table IV.

TABLE IV

Effect on d′ of similarity between test material and interference

	RI	PI	"Location-free" test
Filtering hypothesis	Fall	Fall	No effect
Acid-bath hypothesis	Fall	Fall	Fall
Associative model	Fall	Fall	No effect, or rise

From the results shown in Table III, the only one of these three hypotheses that can be eliminated is the "acid-bath" model, which predicts a fall in d' in the location-free test. The other two hypotheses do not emerge in a much more favourable light, since they both predict a fall in the PI condition, which is not found here (though there is a slight drop in detection rate): both might be modified to predict that confusion arising from a preceding item is less than from a succeeding item, in line with the general finding in short-term memory that proactive

interference is much weaker than retroactive. It would therefore be desirable to devise a further test that would eliminate one of these hypotheses, and this was done in the next experiment. Here, similarity relations were varied only between the pairs of stimuli, and not between the probes and the stimuli: this was achieved by making each stimulus a trigram using one each out of three acoustic classes of letter, "similar" stimuli being ones which used the classes in the same order (thus "rhyming" with each other), while "dissimilar" stimuli were pairs of trigrams which did not rhyme. The probe was again a single letter, and the same three questions were asked about it: but this time, a subject could not be helped in deciding whether the probe had been in the first or second stimulus by recalling the acoustic class of the stimulus—since both stimuli contained three acoustic classes. Thus, he could not use "filtering" in this experiment, and the filtering hypothesis could not be invoked if a fall in d', in RI or PI or both, was found.

Experiment III
Method
Presentation

Items were presented visually in the following order:
      Trigram 1—Trigram 2—Digit Shadowing—Probe Letter + question.
Each of these sections of a trial were separated by a blank frame. Presentation was on a Bina-View display, and the procedure followed was that described in Experiment II.

Stimulus material

The two trigrams both contained one letter from each of the following classes: FLMNSX, BCDEGPTV, AHJK. The trigrams were either "similar" or "dissimilar" to each other: if similar, they had the acoustic classes in the same order, and thus rhymed (e.g. PMJ–BFA); if dissimilar, the representatives of each acoustic class occupied a different position in the two blocks (e.g. JFC–NBK). Half the pairs of trigrams were similar and half dissimilar, the two conditions occurring in random order. No letter was ever presented in both blocks. To eliminate meaningful or familiar trigrams, a list of all possible trigrams was compiled and two of the subjects were asked to select, at leisure, trigrams which "meant anything to them": these trigrams were then censored by the computer program which composed the lists. Apart from these two constraints, the blocks were constructed entirely at random from the available possibilities.

Shadowing

This consisted of digit strings similar to the last experiment.

Probe

Subjects were again presented with a single letter, followed by a symbol indicating one of three questions (was it in the first trigram, the second or in either?). Each question was asked equally often following similar and dissimilar stimuli, and an equal number of new and old probes was employed.

Schedule

Again, several dozen different lists of varying complexity were generated. The same procedure was followed as in Experiment II regarding marking of lists: the subjects were six members of the A.P.U. subject panel, five female and one male, who completed 10 lists after being given one or two for practice.

Analysis

The analysis and presentation of results (Table V) follows that given for the last experiment.

TABLE V

Results of Experiment III (N = 6)

	RI		PI		"Location-free"	
	Sim	Diss	Sim	Diss	Sim	Diss
% Detections	62·2	68·5*	82·7	81·0	78·1	77·3
% False–positives	17·3	17·3	18·8	16·9	17·8	15·3
Log β	−0·92	−0·48	0·02	−0·34	−0·46	−1·18
d'	1·17	1·52*	1·91	1·92	1·91	1·80

Note: "Sim" = similar interference; "Diss" = dissimilar interference; * difference significant at $P < 0.05$ by Wilcoxon matched-pairs signed-ranks test, two-tailed.

Results

Table V shows a considerable reduction of d' with similar RI, but none with PI or in the "location-free" test. The previous experiment eliminated the "acid-bath" hypothesis; now, the "filtering" hypothesis can be eliminated, since the design of the experiment precludes it, leaving only the last of the three explanations considered here, the associative model. (Of course, this does not preclude the possibility that "filtering" was employed under the different conditions of Experiment II.)

It should be noted, however, that the associative model needs some elaboration to explain the effect, given the peculiarities of this experiment. It appears that when one trigram is used as test material and another as interference, a letter in the test trigram is more affected by its positional counterpart in the interference trigram than by the other two constituents of the latter. This suggests that associative confusion is a function of positional factors as well as phonemic similarity, so that the notion of "slotting" and the phonemic-associative model must in fact be combined if either is to account for these results.

References

BROADBENT, D. E. (1971). *Decision and Stress*. London: Academic Press.

BROWN, J. (1958). Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology*, **10**, 12–21.

CLARKE, F. R., BIRDSALL, T. G. and TANNER, W. P. Jr. (1959). Two types of ROC curves and definitions of parameters. *Journal of the Acoustical Society of America*, **31**, 629–30.

INGLEBY, J. D. (1969). Decision-making processes in human memory and perception. Ph.D. Thesis, University of Cambridge, Cambridge.

POSNER, M. I. (1967). Short-term memory systems in human information processing. *Acta Psychologica*, **27**, 267–84.

TANNER, W. P. Jr. and SWETS, J. A. (1954). A decision-making theory of visual detection. *Psychological Review*, **61**, 401–9.

UNDERWOOD, B. J. (1945). The effect of successive interpolations on retroactive and pro-active inhibition. *Psychological Monographs*, **59**, No. 3.

WICKELGREN, W. A. (1965). Short-term memory for phonemically similar lists. *American Journal of Psychology*, **78**, 567–74.

WICKELGREN, W. A. (1966a). Short-term recognition memory for single letters and phonemic similarity of retroactive interference. *Quarterly Journal of Experimental Psychology*, **18**, 55–62.

WICKELGREN, W. A. (1966b). Phonemic similarity and interference in short-term memory for single letters. *Journal of Experimental Psychology*, **71**, 396–404.

THE ACCURACY OF NON-VISUAL AIMING

W. D. A. BEGGS, JACQUELINE A. ANDREW, MARTHA L. BAKER, S. R. DOVE, IRENE FAIRCLOUGH AND C. I. HOWARTH

Department of Psychology, University of Nottingham

A visual feedback theory of movement control developed earlier predicted that the accuracy of movements performed in darkness would be unaffected by movement speed, but be linearly related to distance. We found that for variable errors this is partly true; a linear relationship to distance was demonstrated, but an optimum speed was found. Constant errors showed interesting trends, showing the usual range effect but also being affected by movement time. Finally, a small irreducible component in these errors was found, which was considered to be caused by physiological tremor. These results are easily encompassed in our feedback theory, but would be difficult for an information theory based model of movement control.

Introduction

Movement control by some type of feedback mechanism has been considered by several writers. Such feedback could be kinaesthetic, Gibbs (1954), and Crossman and Goodeve (1963); or visual, Woodworth (1899), Craik (1947*a*, *b*), Vince (1948) and Howarth, Beggs and Bowden (1971).

Basing their theory on the work of Craik, these last workers demonstrated intermittent control of movements (Beggs and Howarth, 1970). They postulated that terminal accuracy, measured from the distribution of hits around the target, is a linear function of the distance from the target at which a final correction is likely to be made; they were able to demonstrate a constant angular error in aiming, which they called σ_θ. There was also a non-reducible tremor component to be considered (Howarth *et al.*, 1971; Beggs and Howarth, 1972). This approach is an alternative to the information theory based views of movement control developed by Fitts (1954) and Welford (1968).

In a situation where the subject aims at a target in the dark, having previously seen it, then the visual feedback theorists would all predict that terminal accuracy depends entirely on initial impulse accuracy, to borrow Woodworth's term. It is difficult to see how the information theory hypothesis can satisfactorily cope with this situation.

We have called this type of task non-visual aiming; we use the term non-visual to avoid confusion with the term ballistic, in its strict sense, as first used by Stetson (1905). Non-visual movements are made at any speed, but cannot be modified by vision during the course of the movement. In terms of our theory, in this

situation, accuracy should be primarily a function of the movement distance, given by

$$\sigma^2 = \sigma^2{}_0 + (\sigma_\theta D)^2$$

where σ^2 is the variance of the error on target, $\sigma^2{}_0$ is a variance attributable to tremor, σ_θ is the angular accuracy of aiming, and D is movement distance.

This function is analogous to equation (2) in Howarth *et al.* (1971), where D was replaced by d_w, the distance away from the target at which the final correction was made.

Both Woodworth (1899) and Vince (1948) asked subjects to close their eyes at the beginning of the movement, and open them at the end. They reported that, for repetitive movements, movement speed had virtually no effect on terminal accuracy, although Keele and Posner (1968) found that accuracy decreased when movement speed increased.

Brown, Knauft & Rosenbaum (1948) studied non-visual aiming in two directions and two planes and distances from 0·6 to 40 cm. They broke their measurement of accuracy into constant and variable errors. Variable errors increased with distance in an almost linear fashion; movements made away from the body were more variable than movements made towards the body. They attributed this to the positions of the terminal points of the movements with respect to the body.

Constant errors were mainly affected by movement distance, and showed the usual range effect; that is, short movement distances produced overshoots, which became undershoots as distance increased. This phenomenon is well documented (Fitts, 1951).

Begbie (1959) found that both direction and distance to the target affected accuracy, under conditions of non-visual aiming. In his pencil and paper experiments aiming accuracy was measured by the mean distance of the subjects' pencil strokes from the target: no measure of distribution was taken. He found that these sorts of error were not linearly related to movement distance.

Beggs (1971) obtained the expected linear relationship between movement distance and variable error, for two practised subjects, although there appeared to be some effect of movement speed on these sorts of error: there appeared to be an optimum rate of about 800 ms.

Values of σ_θ and σ_0 were obtained, when the data were fitted to the equation above; these were respectively 38' of angle and 3·3 mm. These values are similar to those obtained by Howarth *et al.* (1971) and Beggs and Howarth (1972).

In this paper we test our assumption that non-visual aiming accuracy is linearly related to distance, and independent of speed of movement.

Method

Materials

Subjects, right-handed only, sat alongside a table. Clamped to the table edge beside them was a home base, and at a variable distance from it was a horizontal target board, also clamped to the table, on which a graph paper target was held by strip magnets. The subjects' task was to hit the centre of a cross drawn on the graph paper, with one arm of the cross in line with the movement. Errors measured relative to each arm of the cross

were therefore either parallel or perpendicular to movement direction; in Fitts' (1954) experiment, errors would have been measured parallel to movement direction only.

We measured mean square error about the two arms of the cross and mean error. From these two measures, a true constant and variable error score is available.

The set-up differed from Fitts' (1954) experiment in that (a) one target only was used; (b) errors were measured in two dimensions; (c) the movement was a reaching one rather than a swinging one.

The home base and target board were hinged, and had a micro-switch underneath each. These controlled a simple electro-mechanical bistable, which in turn controlled the only source of illumination, an Anglepoise lamp. The result of this was that leaving home base switched off the light, and touching the target switched it back on again. Clearly these conditions constituted non-visual aiming, when the subject aimed repetitively, paced by a metronome.

Procedure

Ten movement speeds were used: 40, 50, 60, 75, 85, 100, 120, 133, 150 and 171 ticks/min on the metronome, being almost exactly 1500, 1200, 1000, 800, 700, 600, 500, 450, 400 and 350 ms movement times, respectively.

Five movement distances were used, being 10, 20, 30, 40 and 50 cm between the centre of the target cross, and an aluminium angle stop on the home base.

At each distance, six naive subjects were used, right-handed, with normal, or corrected-to-normal vision, being undergraduate male and female students. Subjects were asked to tap the target and home base alternately with a pencil in time to the ticks of the metronome; they made 20 hits on each target.

The order of presentation of speeds was randomized three times; subjects were paired, and so the design was balanced for practice and fatigue effects. The same randomizations were used for each group of six subjects.

Subjects were instructed to aim accurately for the centre of the cross, while keeping in time with the metronome; they were given a few trials at various speeds to familiarize them with the experiment.

Results

Targets were scored for mean square accuracy and for mean accuracy. The former is a measure combining both the mean and variance, or constant and variable error scores, given by

$$M.S. = M^2 + \sigma^2.$$

To find the true variance of the distribution, this measure and the mean are required. In accordance with on-going work in this area, the following method of signing the constant errors was adopted.

Clearly, errors measured parallel to the main movement direction can either be overshoots or undershoots, and can be signed positive and negative respectively. In the case of errors measured perpendicular to the movement direction, the situation is not so clear. We have decided that in this case, errors which are on the left of the target line are undershoots, and scored negative, and vice versa.

This convention considers errors relative to the subject's egocentre, and not his hand starting position. We feel our data are more orderly when treated in this way.

From the mean square and mean values of the distribution, measured separately in the two directions parallel and perpendicular to movement, we were able to

derive mean and variance scores for each subject at each distance, for each speed, measured parallel and perpendicular to movement direction.

The design of this experiment can be considered as the Three Factor with Repeated Measures (Case 1) type discussed by Winer (1962) pp. 319 *et seq.*

Analyses of variance on the variable and constant error scores were run and produced the results summarized in Tables I and II.

TABLE I

The results of an analysis of variance on the variance of the error scores for six naive subjects

Source	SS	df	VE	F	P
Between subjects					
A (distance)	380881	4	95220	22·46	<0·001
Subject within groups	1282·46	25	51·30		
Within subjects					
B (speed)	24735·37	9	2748·37	2·22	<0·05
A × B	49688·32	36	1380·23	1·11	NS
B × subjects within groups	278784·07	225	1239·04		
C (error measurement)	20072·55	1	20072·55	21·12	<0·001
A × C	17666·76	4	4416·69	4·65	<0·01
C × subjects within groups	23758·12	25	950·32		
B × C	2767·03	9	307·45	0·42	NS
A × B × C	23206·64	36	644·63	0·89	NS
BC × subjects within groups	163166·02	225	725·18		

TABLE II

The results of an analysis of variance on the signed mean error scores for six naive subjects

Source	SS	df	VE	F	P
Between subjects					
A (distance)	326·00	4	81·50	1·59	NS
Subjects within groups	1282·46	25	51·30		
Within subjects					
B (speed)	173·68	9	19·30	1·79	NS
A × B	315·85	36	8·77	0·81	NS
B × subjects within groups	2429·96	225	10·80		
C (error measurement)	0·04	1	0·04	0·00	NS
A × C	127·14	4	31·78	0·92	NS
C × subjects within groups	860·95	25	34·44		
B × C	161·49	9	17·94	1·94	NS
A × B × C	381·19	36	10·59	1·15	NS
BC × subjects within groups	2075·91	225	9·23		

Discussion

Variable Errors

The results of the analysis of variance on the variance scores show three things. First, that the effect of distance is very important for these data. Second, that errors measured parallel and perpendicular to movement direction are different in magnitude, and are differently affected by movement distance. Third, that movement speed does have a small effect on these data.

Effect of Distance

Figure 1 shows the relationship of variable errors to distance for the two measurement directions. Two straight lines were fitted to these data. These were

$$\sigma^2 = 11 \cdot 52 + 0 \cdot 00032 \, D^2 \text{ (Parallel)}$$
$$\sigma^2 = 7 \cdot 38 + 0 \cdot 00025 \, D^2 \text{ (Perpendicular)}$$

accounting for 89·60 and 99·57% of the variance respectively.

The slope constants in the above equations are significantly different ($P < 0 \cdot 02$). For errors measured parallel to movement direction, σ_θ is 1°, while, perpendicular to motion, accuracy rises to 54' of angle. These are higher than those found for naive subjects aiming in visual conditions (Howarth *et al.*, 1971), or for practised subjects in non-visual conditions (Beggs, 1971).

This may reflect the difficulty our naive subjects had with the non-visual situation, or be due to the different kinds of movement made in this situation, and on the apparatus used in these earlier studies.
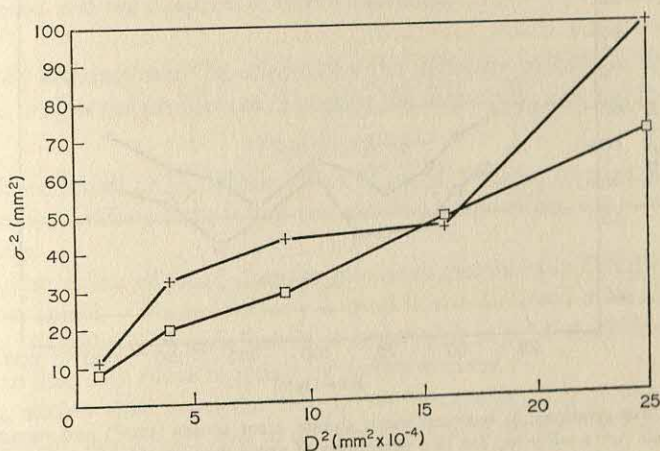
FIGURE 1. The relationship between mean square error scores (mm²) and movement distance (mm² × 10⁻) for six naive subjects, and two directions of error measurement. ×——× = parallel, □——□ = perpendicular.

The intercepts represent the tremor component associated with the movement, measured in two dimensions. They take the form of a variance added to the total variable error. In the direction of movement, tremor with a standard deviation of 3·4 mm was found, while, perpendicular to movement, this dropped to 2·7 mm.

These values of σ_θ are of the same order as those found by Howarth *et al.*, Beggs and Howarth (1972) and Beggs (1971).

The results confirm the hypothesis that the accuracy of non-visual aiming is primarily a linear function of movement distance, and are in agreement with the feedback theory developed by Howarth *et al.* They would be difficult for an information-based theory of movement control, such as Fitts, (1954).

Welford (1968) developed a modification to Fitts' Law, when he considered the movement to be in two parts, a faster distance-covering phase, and a homing phase. The resulting prediction of terminal accuracy contained a term specifying

ballistic accuracy, or non-visual accuracy. In fitting the results of Welford, Norris and Shock (1963) to this function he assumed that this term was independent of movement distance. Our data show that variable error is a linear function are of movement distance, and therefore the data is less well fitted by his equations than he believed.

Effect of Rate

Figure 2 shows that variable errors measured parallel to movements are higher than errors measured perpendicular to movement. This implies an elliptical distribution of hits, which is confirmed by the earlier result that σ_θ differed significantly for the two error measurement directions. Such a result had been found earlier by Crossman (unpublished; reported in Crossman and Goodeve, 1963).
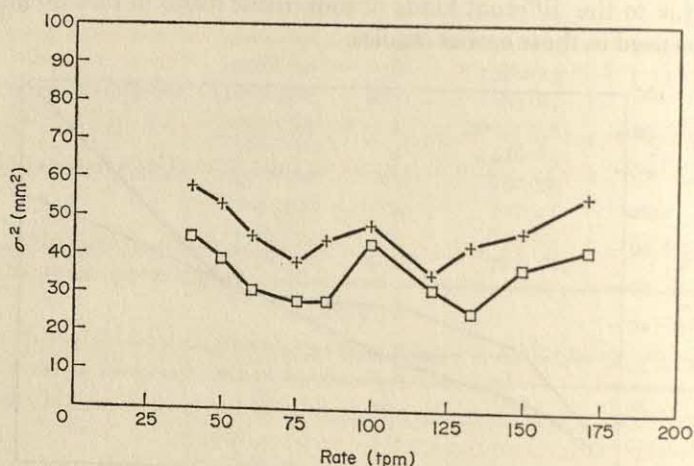
FIGURE 2. The relationship between mean square error scores (mm²) and movement rate in ticks/min for six naive subjects, for two directions of error measurement. ×——× = parallel; □——□ = perpendicular.

The second point arising from these data is the nature of the curve relating terminal error to rate. This is essentially U-shaped, although two individual values of σ^2 distort the curve at a speed of 100 tpm. In these experiments, one very large error in aiming can distort the variance of the distribution disproportionately; we believe a U-shaped curve to be a good description of our data.

Woodworth's data showed this, for both hands, although more pronounced for the left hand over the range of speeds which we used.

Vince's results showed an increase in error at lower speeds, but as she did not use speeds above 100 tpm no corresponding increase at faster speeds was noted.

Keele and Posner (1968), using speeds in the high-range, 450–150 ms movement time, found that errors increased with increasing speed.

Beggs found a somewhat similar relationship between variable errors and rate, with an optimum at about 800 ms movement time.

The reasons for this curvilinear effect of rate on error are not known: it may be that undetected tremor is creeping in at slow speeds, while at the faster speeds,
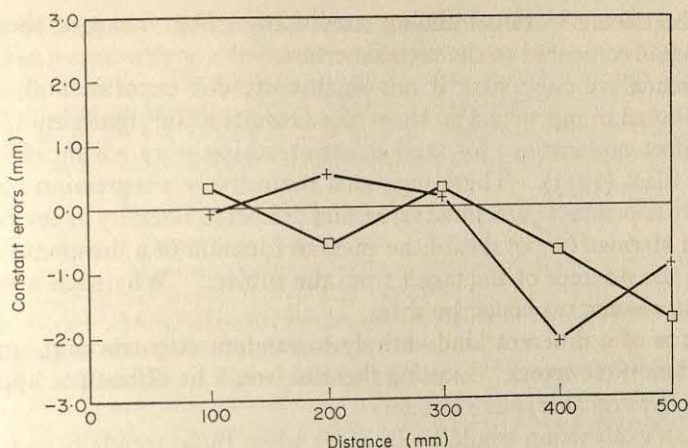
FIGURE 3. The relationship between constant errors (mm) and movement distance in mm for six naive subjects, and two directions of error measurement. ×——× = parallel; □——□ = perpendicular.

initial aiming accuracy may be affected by the necessity to keep in time with the metronome, and/or the problem of relocating the target after a period in the dark.

Constant Errors

Although we found no significant effect of speed, distance or error measurement direction, other workers have found the classical range effect, for visual and non-visual aiming.

Figure 3 shows the effect of distance on errors measured in two directions, collapsed across speed. There is clearly a trend in the direction of increased undershooting as distance increases, but it is impossible to say if this is linear. These results are in line with those reported by earlier workers.

Figure 4 shows that undershooting errors are particularly prone to occur at short movement times, and less so at long movement times. We have evidence
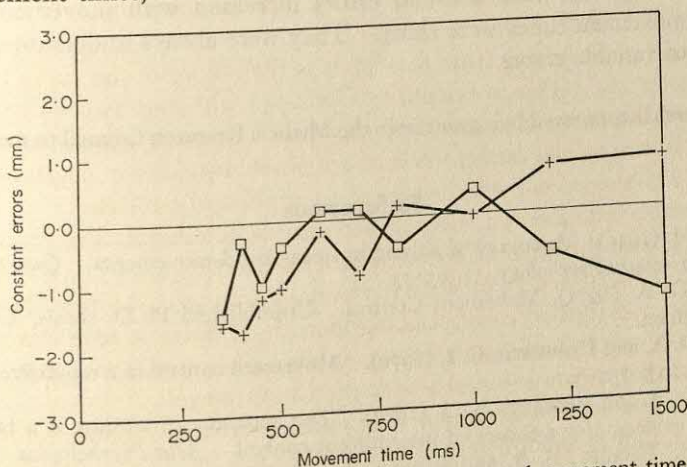
FIGURE 4. The relationship between constant errors (mm) and movement time (ms), for six naive subjects and two directions of error measurement. ×——× = parallel; □——□ = perpendicular.

that this also occurs in visual aiming conditions. The values of these constant errors are small compared to the variable errors.

These trends are suggestive if not significant; our experimental design may have contributed to our failure to show these results to be significant.

Range effect explanations for random step tracking were advanced by Johnson (1952) and Slack (1953). These suggested respectively a regression phenomenon or a memory dependent perceptual error, and predicted linearity of the range effect. Pepper and Herman (1970) regard the effect as function of a "response set" which varies with the distance of the target from the subject. Why such a response set should exist was not explained by them.

Our task is of a different kind entirely to random step tracking, and shows a speed effect on these errors. Existing theories would be difficult to apply to these results: we have no alternative to offer.

Perhaps an explanation would be in order when these trends have been shown to be significant.

Conclusion

Our experiment showed that, as hypothesized, variable errors were a linear function of movement distance, for non-visual aiming.

Naive subjects had an angular accuracy of $1°$ when errors were measured parallel to movement direction; when they were measured perpendicular to movement direction, the angular accuracy rose to $54'$ of angle.

Tremor components of error were demonstrated; the standard deviations of errors due to this were $3·4$ mm parallel to movement, and $2·7$ mm perpendicular to movement direction.

A small effect of rate was shown; there appeared to be an optimum rate at about 500 ms movement time.

We also collected some data on constant errors. While unsignificant, trends were seen which suggested that constant errors increased with movement distance, and when movement times were short. They were always undershots, and small compared to variable errors.

References

BEGBIE, C. H. (1959). Accuracy of aiming in linear hand-movements. *Quarterly Journal of Experimental Psychology*, **11**, 65–75.

BEGGS, W. D. A. (1971). Movement Control. Unpublished Ph.D. thesis, University of Nottingham.

BEGGS, W. D. A. and HOWARTH, C. I. (1970). Movement control in a repetitive motor task. *Nature*, **225**, 752–3.

BEGGS, W. D. A. and HOWARTH, C. I. (1972). The accuracy of aiming at a target—some further evidence for a theory of intermittent control. *Acta Psychologica*. (In press.)

BROWN, J. S., KNAUFT, E. B. and ROSENBAUM, G. (1948). The accuracy of positioning reactions as a function of their direction and extent. *American Journal of Psychology*, **61**, 167–82.

CRAIK, K. J. W. (1947a). Theory of the human operator in control systems: I. The operator as an engineering system. *British Journal of Psychology*, **38**, 56–61.

CRAIK, K. J. W. (1947b). Theory of the human operator in control systems: II. Man as an element in a control system. *British Journal of Psychology*, **38**, 142–8.

CROSSMAN, E. R. F. W. and GOODEVE, P. J. (1963). Feedback control of hand-movement and Fitts' Law. Unpublished. Communication to Experimental Psychology Society, Oxford.

FITTS, P. M. (1951). Engineering psychology and equipment design. In STEVENS, S. S. (Ed.), *Handbook of Experimental Psychology*. New York: Wiley.

FITTS, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, **47**, 381–91.

GIBBS, C. B. (1954). The continuous regulation of skilled response by kinaesthetic feedback. *British Journal of Psychology*, **45**, 24–39.

HOWARTH, C. I., BEGGS, W. D. A. and BOWDEN, J. (1971). The relationship between speed and accuracy of movement aimed at a target. *Acta Psychologica*, **35**, 207–18.

JOHNSON, D. M. (1952). The central tendency of judgement as a regression phenomenon. *American Psychologist*, **7**, 281.

KEELE, S. W. and POSNER, M. I. (1968). Processing of visual feedback in rapid movements. *Journal of Experimental Psychology*, **77**, 155–8.

PEPPER, R. L. and HERMAN, L. M. (1970). Decay and interference effects in the short-term retention of a discrete motor act. *Journal of Experimental Psychology*, **83**, Monograph Supplement.

SLACK, C. W. (1953). Perception and past experience: a mathematical model for Taylor's "range effect". *American Psychologist*, **8**, 437.

STETSON, R. H. (1905). A motor theory of rhythm and discrete succession. *Psychological Review*, **12**, 250–70.

VINCE, M. A. (1948). Corrective movements in a pursuit task. *Quarterly Journal of Experimental Psychology*, **1**, 85–106.

WELFORD, A. T. (1968). *Fundamentals of Skill*. London: Methuen.

WELFORD, A. T., NORRIS, A. H. and SHOCK, N. W. (1963). Movement, time and range; a preliminary report (abstract). *Ergonomics*, **6**, 310.

WINER, B. J. (1962). *Statistical Principles in Experimental Design*. London: McGraw Hill.

WOODWORTH, R. S. (1899). The accuracy of voluntary movement. *Psychological Review*, Monograph Supplement, Vol. III, No. 2 (Whole No. 13).

DISRUPTION OF FREE-OPERANT SUCCESSIVE DISCRIMINATIONS BY SEPTAL DAMAGE IN RATS

ANTHONY DICKINSON

Laboratory of Experimental Psychology, University of Sussex

In Experiment I rats were trained preoperatively on a successive, free-operant discrimination with fixed component durations until S— responding had been minimized. Septal damage resulted in a transitory loss of discriminative performance due to an elevation of S— responding. Experiment II showed a comparable septal deficit following training which employed a "correction" procedure that penalized S— responding. It is concluded that, even after extended training, the maintenance of a low level of S— responding by rats in these types of discriminations requires the operation of an active mechanism which can be disrupted by septal damage. The possibility that this disruption results from a lesion-induced change either in the primary reactions to frustrative non-reward or in associated processes is discussed.

Introduction

Several authors (Thomas, Hostetter and Barker, 1968; Caplan, 1970; Gray, 1970) have suggested that septal damage in the rat alters the primary reactions to frustrative non-reward. The empirical basis for this idea comes mainly from the behavioural effects of septal lesions on performance in extinction following continuous and partial reinforcement (Gray, 1970) and on response patterning under schedules of positive reinforcement which require a low rate of responding (Thomas *et al.*, 1968; Caplan, 1970). However, this explanation can be extended to cover other behavioural changes demonstrated by septal animals. Schwartzbaum, Kellicut, Speith and Thompson (1964) trained two groups of rats on a successive, free-operant discrimination before administering septal lesions to one group. During the initial post-operative sessions, the septal rats showed almost total loss of discriminative performance due to an elevation of responding in the non-rewarded stimulus (S—). The control subjects, on the other hand, demonstrated perfect transfer of the discrimination across the surgical procedure and post-operative recovery period. As there is good evidence that emotional behaviour occurs in the S— stimulus following the withdrawal of reward (Terrace, 1966a; Daly, 1971) and that such behaviour plays a causal role in mediating the response decrement, the elevated responding of septal rats in this stimulus could result from a lesion-induced change in the reaction to frustrative non-reward.

Amsel and Ward (1965) have suggested that primary reactions to frustrative non-reward only play a temporary role in the maintenance of a successive discrimination. They showed that the "frustration effect" following non-reward

on S— trials disappears as asymptotic discrimination is attained.   According to frustration theory (Amsel, 1958), the necessary antecedent condition for non-reward to be frustrating is that the animal should expect reward in the presence of the accompanying stimuli, and Amsel and Ward (1965) proposed that the decline in the "frustration effect" resulted from a concurrent decrease in the reward expectancy elicited by the S— as the successive discrimination was formed. Consequently, if septal damage affects behaviour following reward withdrawal by altering the magnitude of the reactions to frustrative non-reward, the lesion should only exert a disruptive effect on the maintenance of a successive discrimina-tion for as long as the animal continues to expect reward in the presence of the S— stimulus during preoperative training.   As the mean terminal pre-operative response rates in the S+ and S— stimuli of the Schwartzbaum et al. (1964) experiment were 45·5 and 12·2 response per min respectively, there are good grounds for arguing that these subjects maintained a significant S— reward expect-ancy at the time of surgery.   In Experiment I rats were given fairly extensive pre-operative training on a successive, free-operant discrimination to investigate whether septal damage would produce a disruptive effect when S— responding had been almost totally eliminated.

Experiment I

Method

Subjects

The subjects were 12 experimentally naive male hooded Lister rats with a mean free-feeding weight of 346 g (range 284–425 g) at the start of the experiment.   They were gradually reduced to 85% of their free-feeding weight before the start of training and were then maintained at this level throughout the experiment, except during the post-operative recovery period, by being given a restricted amount of food after each daily session.

Apparatus

Subjects were tested in an operant-conditioning chamber housed in a sound-attenuating cubicle supplied with a white noise masking source.   The chamber (20 × 20 × 20 cm) was painted matt black on the inside and had a 24-V 3-W houselight mounted in the roof.   The voltage supplying this houselight was dropped to 12–14 V during experimental sessions.   A single Campden Instrument Rat Lever was mounted on the front wall at a height of 4 cm above the floor level and offset to one side of the centre point by 4 cm.   A recessed magazine, which could be illuminated by a 24-V 3-W bulb, was set in the front wall adjacent to the lever.   A heavy duty relay was mounted on the outside of the front wall above the lever to provide auditory feedback for lever presses.   A 21-W stimulus light was also mounted on the front panel 12·5 cm above the lever.   Reinforcement consisted of a single 54–mg Dixon's diet 41 b pellet.   The apparatus was controlled by conventional electromechanical pro-gramming equipment.

Surgery

At the time of surgery the septal subjects were anaesthetized by pentobarbital (35–40 mg/kg) with ether supplement.   Bilateral lesions of the septal nuclei were produced electro-lytically by passing a 2-mA d.c. anodal current for 20 s through the 0·5-mm uninsulated tip of a stainless steel needle.   The frame of the stereotaxic instrument acted as the indifferent electrode.   The electrode was placed stereotaxically with the skull horizontal 1·0 mm anterior

to the bregma, 3·0 mm lateral to the midline, and 6·0 mm below the surface of the skull and angled 25° towards the midline. Control rats were subjected to a surgical procedure identical to that of the septal animals except that the electrode was not lowered into the brain.

All subjects were given 10 days post-operative recovery during the first six of which they had free access to food and water before being returned to their 85% free-feeding weight.

Procedure

After magazine training, six animals were shaped to press the lever and given 60 reinforcements on a CRF schedule. The delivery of reinforcement was accompanied by the illumination of the magazine for 3 s. Responses emitted during this reinforcement period were not recorded and had no programmed consequences. Over the next two sessions 60 reinforcements were given on a VI-15 s and on a VI-30 s schedule respectively before the subjects were introduced to pre-operative baseline training.

Daily baseline sessions consisted of the successive presentation of 24 components, each of 120 s duration and separated from the next component by a 10-s intercomponent interval. The stimulus light was on during half of the components (light-on) and off during the remaining components (light-off). The houselight was on continuously except during the intercomponent intervals. Each session started with two alternations of the light-on and light-off components followed by two series each consisting of 5 light-on and 5 light-off components. Within each of the two series, the order of presentation of the components was random with the restriction that neither type of component could occur more than three times in succession. The order of presentation of the two series was alternated across daily sessions. Responding in both components was reinforced on a VI-45 s schedule. The reinforcement schedules operative in the two types of component were independent. Reinforcement "set up" but not collected in a component was cancelled at the end of that component. Responses emitted during the inter-component interval had no programmed consequences. Each response produced auditory feedback except during the inter-component interval and the 3-s reinforcement period.

The subjects were given 30 pre-operative baseline sessions under these conditions before the introduction of discrimination training. The subjects received 30 sessions of pre-operative discrimination training before septal lesions were administered. For half the subjects, responding in light-on components was extinguished whilst the schedule of reinforcement for responding in light-off components remained VI-45 s. For the remaining subjects, the opposite relationship was established. Components in which responding was extinguished will be referred to as S— components, and those in which responding was reinforced will be referred to as S+ components. Following post-operative recovery, a further 18 sessions of discrimination training were given.

Six further rats were exposed to a procedure identical to that employed for the septal animals except that they were all given sham operations after 30 sessions of pre-operative discrimination. These animals will be referred to as the control group.

Results

Histology and Affective Changes

All the lesions destroyed most of the pre-commissural septum from the genu of the corpus collosum to the crossing of the anterior commissure. The nucleus and tract of the diagonal band of Broca and the dorso-medial tip of the nucleus accumbens also usually sustained some damage. Typically, the lesion extended beyond the pre-commissural septum, involving the hippocampus pars anterior and the medial parolfactory area in the rostral direction and the supra-commissural region of the lateral nucleus, the nucleus septalis fimbralis, the nucleus triangularis septi, the descending columns of the fornix, and the fornix itself in the caudal direction. Occasionally there was also some damage to the medial portion of the putamen.

The anterior commissure and the preoptic areas remained intact, and the corpus callosum and overlying cortex only received slight damage from the electrode tract.

All of the septal rats showed some signs of the classic hyper-irritability syndrome (Brady and Nauta, 1953) during the first few post-operative days. This syndrome was much attenuated by the end of the recovery period.

Behavioural Results

One septal rat failed to emit any responses in either component during the first two post-operative sessions, and one control rat died during the sham operation. The results of these two animals have been excluded from this presentation.

There was no significant difference between the response rates of the two groups in S— components during the pre-operative discrimination sessions ($F < 1$) and no significant Group X Session interaction ($F < 1$). However, the control group showed a higher response rate than the septal group in S+ components during the pre-operative discrimination sessions ($F = 6.25$, $df = 1, 8$, $P < 0.05$).

The response rates of the two groups in S+ and S— components during the last five pre-operative discrimination sessions (sessions 26–30) and during all the post-operative sessions (sessions 31–48) are illustrated in Figure 1. First, it can be seen that the pre-operative discrimination training succeeded in establishing a very low rate of responding in S— components during the sessions immediately preceding surgery. Second, Figure 1 shows that the septal animals exhibited a marked increase in S— responding during the initial post-operative sessions above
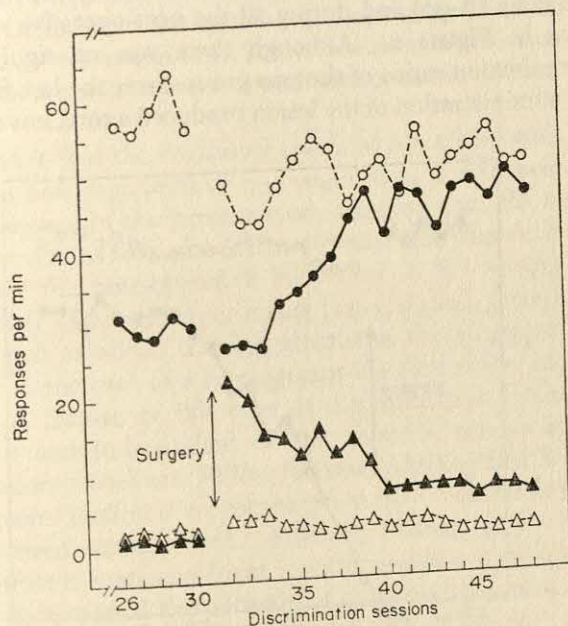
FIGURE 1. Mean response rates of the septal and control groups in S+ and S— components during the last five pre-operative (sessions 26–30) and all the post-operative discrimination sessions (sessions 31–48). ●——● S+ septal; ▲——▲ S— septal; O————O S+ control; △————△ S— control.

their pre-operative rate and the post-operative rate of the control group. With
further training the S— response rate of the septal animals declined to an asymp-
totic value which was still above the rate of the control group. The surgical pro-
cedure and post-operative recovery period had little effect on the S— response
rate of the control group. Because of the low values of the post-operative S— rate
for the control group, the transformation $X' = \sqrt{(X + 0.5)}$ was performed on
the raw data before statistical analysis. There was a significant difference between
the post-operative S— response rates of the two groups ($F = 28.16$, $df = 1, 8$,
$P < 0.005$) and a significant Group X Session interaction ($F = 2.31$, $df = 17$,
136, $P < 0.01$).

The response rate of the septal group in S+ components during the initial
post-operative sessions was similar to the terminal pre-operative S+ rate. With
further training the septal animals showed a progressive increase in S+ rate, whilst
the S+ rate of the control group remained relatively stable across post-operative
sessions. During these post-operative sessions there was no significant effect of
the lesion on S+ response rate ($F < 1$), but the Group X Session interaction
did reach significance ($F = 2.76$, $df = 17, 136$, $P < 0.005$).

As the pre- and post-operative S+ rates of the two groups differed, the effect
of the lesion on a relative measure of discriminative performance was assessed.
For this purpose a discrimination ratio was calculated for each subject on each
session. This ratio equalled the total number of responses emitted in S+
components divided by the sum of the total number of responses emitted in both
S+ and S— components. The mean ratios of the two groups over the last five
pre-operative (sessions 26–30) and during all the post-operative sessions (sessions
31–48) are shown in Figure 2. Although there was no significant difference
between the discrimination ratios of the two groups over the last five pre-operative
sessions ($F < 1$), administration of the lesion produced a total loss of discriminative
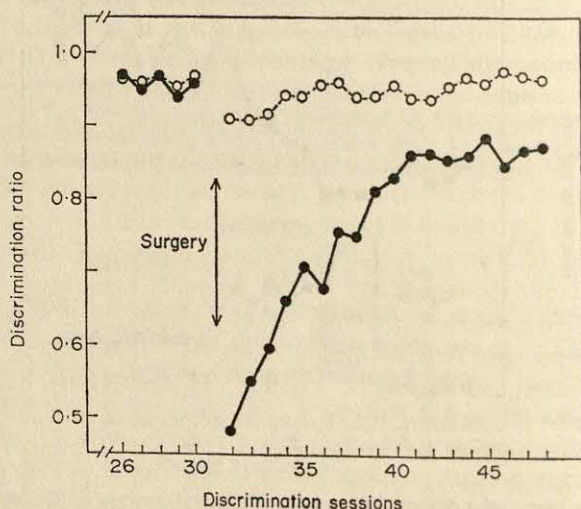
FIGURE 2. Mean discrimination ratios of the septal and control groups during the last five
pre-operative (sessions 26–30) and all the post-operative discrimination sessions (sessions 31–48).
●——● septal; O————O control.

performance on the first post-operative session followed by a progressive re-acquisition of discrimination to a lower asymptote than the control group. There was a significant effect on the lesion ($F = 24.60$, $df = 1$, 8, $P < 0.005$) and a significant Group X Session interaction ($F = 6.90$, $df = 17$, 136, $P < 0.005$) on the post-operative discrimination rate. Before statistical analysis the transformation $X' = 2 \arcsin \sqrt{X}$ was performed on the discrimination ratios.

Discussion

Whilst the control subjects maintained the discrimination following the surgical and post-operative recovery period, the septal group showed a marked increase in S— responding even though the pre-operative response output in S— components was minimized by fairly extensive training. As the subject selection procedure unfortunately resulted in two groups whose pre-operative S+ rates differed significantly, this increase cannot be unequivocally ascribed to septal damage. However, it is unlikely that the group difference in post-operative S— responding was related to the difference in pre-operative S+ rate. In order to explain the results in terms of this difference, one would have to assume that an inverse relationship existed between the pre-operative S+ response rate and the post-operative S— rate. In fact the S+ rate on the last pre-operative session and the S— rate on the first post-operative session were positively correlated for the septal group ($r = 0.85$) and not correlated for the control group ($r = -0.14$). In view of this finding and the similarity of the terminal pre-operative discrimination ratios and S— response rates of the two groups, it seems reasonable to assume that the loss of discriminative performance by the septal group resulted from infliction of the lesion.

On the hypothesis that the emotional reactions associated with S— components disappear when a near zero level of S— responding is established, the disruptive effect of septal damage in the present study suggests that the lesion does not act by changing the reactions to frustrative non-reward and that an alternative mechanism must underly the post-operative increment in S— output. However, the results of a number of recent experiments provide good grounds for questioning the assumption that extended training attenuates the emotional consequences of S— non-reward in the case of a free-operant discrimination such as employed in Experiment I. A feature of this type of discrimination is that the component duration is fixed and independent of the animal's behaviour. Using a fixed component-duration procedure, Rilling, Askew, Ahlskog and Kramer (1969) have shown that pigeons continue to escape from S— components long after S— responding has been extinguished. Further, Hearst (1971) has reported that behavioural contrast is sustained for at least 64 sessions of discrimination training. Several authors have related the presence of behavioural contrast to occurrence of emotional reactions in the nonrewarded stimulus (Terrace, 1966a; Bloomfield, 1967). Consequently, it is possible that with fixed component-duration training the occurrence of emotional reactions associated with reward withdrawal still plays a role in controlling response output even when this output has been minimized by

extensive training.   It follows that the elevation of S— responding following septal damage could be due to interference with this mechanism via a change in the primary reactions to frustrative non-reward.

Experiment II

The strongest evidence for a decline in the frustrative effects of reward withdrawal with overtraining on a free-operant discrimination comes from studies utilizing a "correction" procedure.   Under this procedure responding in an S— component delays the termination of the current component.   Terrace has demonstrated a decline in contrast, "peak shift", and escape from the S— stimulus when extended discrimination training is given with a "correction" procedure (Terrace, 1966b, 1971).   Consequently, it is possible that septal damage would not have a disruptive effect on discrimination performance if extended training with a "correction" procedure was given prior to surgery.   Experiment II was conducted to investigate this possibility.

Method

The subjects were six experimentally naive male hooded Lister rats with a mean free-feeding weight of 295 g (range 276–320 g) at the start of the experiment.   The apparatus, deprivation regime, and surgical and post-operative recovery procedure were identical to those used in Experiment I.   The subjects were magazine trained, shaped to press the lever, and trained to respond on a VI-15 s and VI-30 s schedule following the procedure used in Experiment I.   After this pretraining, 20 baseline sessions were administered.   As in Experiment I, baseline sessions consisted of the successive presentation of 12 light-on and 12 light-off components.   The component stimuli, duration, and sequencing, and the intercomponent interval conditions were also identical to those employed in Experiment I.   Reinforcements were delivered on a VI-45 s schedule in both components.

Pre-operative discrimination training

Following baseline training the subjects were randomly divided into two groups of three rats each, and given 50 sessions during which responding was extinguished in S— components.   For two subjects in each group, S— components were those accompanied by the illumination of the stimulus and houselights (light-on), and for the remaining subjects, were those accompanied by the illumination of the houselight alone (light-off).   The schedule of reinforcement for responding in S+ components remained VI-45 s.

During pre-operative discrimination training a "correction" procedure was employed.   A response emitted during an S— component stopped the timer determining the duration of that component and started a timer determining the duration of a penalty period.   A response emitted during this penalty reset and restarted the penalty timer.   When a penalty period had elapsed, the component timer was restarted.   The component time accumulated prior to the institution of a penalty period was carried over the penalty period.   No programmed stimulus changes were associated with the institution and termination of a penalty period.   Successive components were still separated by a 10-s intercomponent interval.   If a subject made no S— responses the total programmed duration of a session was 52 min.   A maximum limit of 60 min was set on the session duration.   If an S— component was in force when this maximum limit expired, the session was terminated immediately.   However, if an S+ component was in force when the maximum limit expired, the session was terminated when this current component ended.   The value of the penalty period was 60 s for the first 22 sessions and 120 s for the remaining 28 sessions.   The value of the penalty period was increased to further minimize S— responding.

After pre-operative discrimination training one group of subjects was given bilateral septal lesions whilst the subjects in the other group received sham operations.

Post-operative discrimination training

Following the post-operative recovery period the subjects were given 20 discrimination sessions using a fixed component-duration procedure. The conditions were identical to those employed preoperatively except that responses emitted during S— components did not institute penalty periods.

Results

The size and nature of the septal damage was essentially similar to that produced in Experiment I.

Pre-operative training

Figure 3 shows the mean response rates of the two groups during the last five baseline sessions (points marked B on the abscissa) and during the pre-operative discrimination sessions when a 60-s (sessions 1–22) and a 120-s penalty period (sessions 23–50) were in force. By the end of baseline training the response rates of all the subjects were stable. Figure 3 shows that the introduction of discrimina-
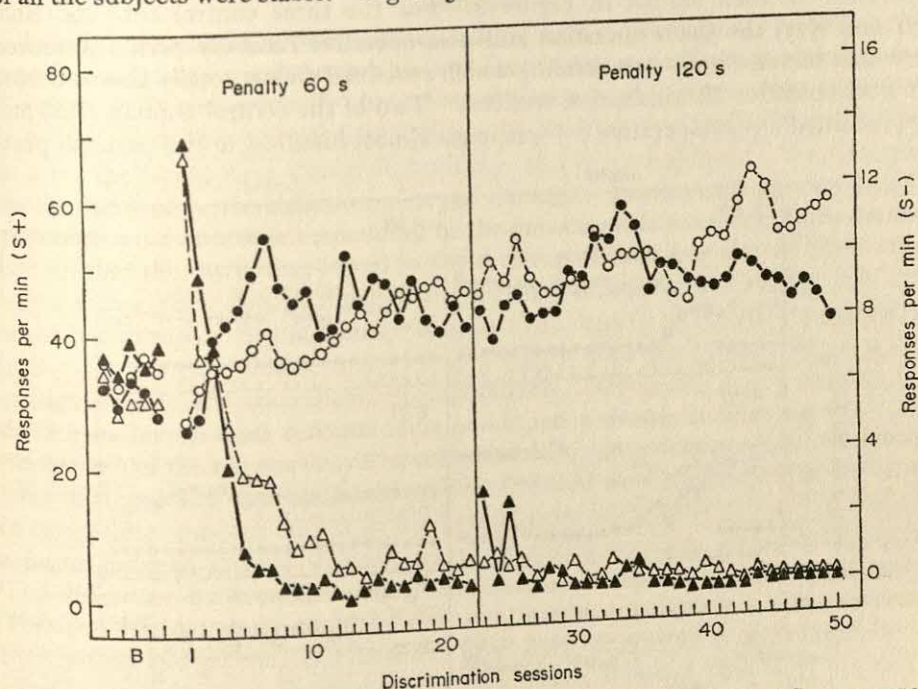
FIGURE 3. Mean response rates of the septal and control groups in S+ and S— components during the pre-operative discrimination sessions when a 60-s (sessions 1–22) and 120-s (sessions 23–50) penalty period was in force. The values plotted against the points marked B on the abscissa represent the mean response rates for each group in S+ and S— components during the last five baseline sessions. The response rate scale for responding in both components during the baseline sessions and in S+ components during the discrimination sessions is represented on the left-hand ordinate, and the scale for responding in S— components during the discrimination sessions on the right-hand ordinate. ●——● S+ septal; ▲——▲ S— septal; O— – – –O S+ control; △– – – –△ S— control.

tion training produced an increase in the mean S+ response rate indicating the occurrence of behavioural contrast. This pattern was typical of individual subjects' performance. Except for one session of one subject, the S+ response rate of each subject during all of the last 10 pre-operative discrimination sessions was higher on every session than the highest rate emitted in the same component during the last five baseline sessions. Figure 3 also shows that there was no general tendency for the magnitude of contrast to decrease with discrimination training.

In order to establish the pre-operative comparability of the two groups a statistical comparison of the performance of the two groups over the last 20 pre-operative sessions was made. There was no significant difference between the S+ ($F < 1$) or the S— response rates ($F = 2\cdot45$, $df = 1, 4$, $P > 0\cdot10$) of the two groups. The transformation $X' = \sqrt{(X + 0\cdot5)}$ was performed on S— responses rates before analysis.

Post-operative training

The response rates in S+ and S— components during the last five pre-operative sessions (sessions 46–50) and during all the post-operative sessions are presented separately for each animal in Figure 4. For the three control subjects (R68, R71 and R75) the sham operation and post-operative recovery period produced very little increase in S— responding which remained almost totally absent during the post-operative discrimination sessions. Two of the control animals (R68 and R71) emitted a post-operative S+ response almost identical to the terminal post-
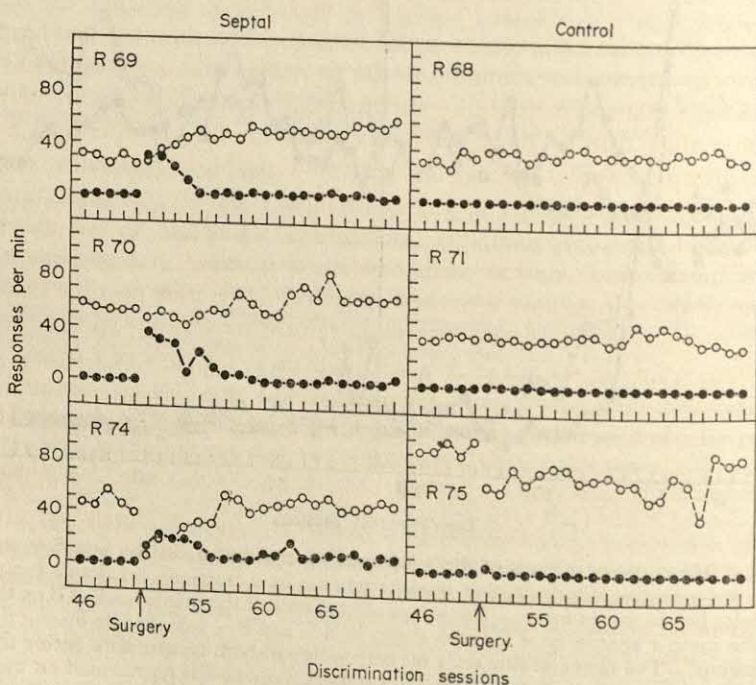
FIGURE 4. Response rates of individual septal (left-hand panels) and control (right-hand panels) subjects in S+ and S— components during the last five pre-operative (sessions 46–50) and all the post-operative discrimination sessions (sessions 51–70). ●——● S—; ○----○ S+.

operative value, whilst for the third control subject (R75) the post-operative rate was lower. For all the control subjects the S+ response rate remained relatively constant across the post-operative sessions.

By comparison to controls, all three septal rats (R69, R75 and R74) showed a substantial increase in responding in S— components during the initial post-operative sessions. On the first post-operative session there was little differentiation between the response output in S+ and S— components indicating that the lesion produced almost complete loss of discriminative performance. For R69 and R70 the S— response rate progressively declined until it reached a level similar to the pre-operative and control response rates. However, the third septal rat (R74) showed a sustained elevation of S— responding. The initial post-operative response rates of R69 and R70 in S+ components were similar to the terminal pre-operative rates, and with further training the S+ rates of these two animals exhibited a progressive increase. The third septal subject (R74) at first showed a decrement in S+ responding following surgery which then increased to a level equivalent to that maintained during the terminal pre-operative sessions.

Discussion

In spite of extensive pre-operative training with a "correction" procedure, the septal rats showed a substantial loss of discriminative performance which, as in Experiment I, was primarily due to an elevation of S— responding during the immediate post-operative sessions. Since the pre-operative rates of the two groups in this experiment were similar in both S+ and S— components, the disruption can be ascribed to the infliction of septal damage. Further, the similarity of the pattern of post-operative responding in the two experiments strengthens the conclusion that the disruption found in the first experiment was also an effect of the lesion rather than due to the discrepancy between the pre-operative S+ rates of the two groups. Septal damage appears to be capable of reinstating responding in S— components almost to its initial strength even after responding has been extinguished in the component pre-operatively. However, in the absence of a decline in behavioural contrast with continued training in Experiment II, the mechanism of the reinstatement remains obscure. It is fully possible that primary emotional reactions to the non-occurrence of reward were still playing an active role in controlling response output in S— components at the time of surgery, and that a lesion-induced change in such reactions would produce an elevated output. The difference between the pattern of contrast found in Experiment II and in Terrace's (1966b) study could be due to a number of factors: species difference between rat and pigeon, differences in procedure, or simply the administration of insufficient training in Experiment II. However, this last possibility appears to be unlikely in view of the fact that two of the control subjects (R68 and R71) maintained the terminal pre-operative S+ response rates across the further 20 sessions of post-operative training.

The disruption of discriminative performance cannot be due to a general memory loss for pre-operative training as, except for R74 in Experiment II, the septal subjects in both experiments showed S+ response rates in the initial post-operative sessions that were very similar to the terminal pre-operative rates. However, a

second possibility is that the lesion produced a specific retention deficit. The selective effect of the lesion could have resulted from a memory loss for response patterns specifically related to stimuli associated with frustrative non-reward. Frustration theory (e.g. Amsel, 1972) has usually assumed that the response decrement observed following the withdrawal of reward is, at least partially, due to the occurrence of competing instrumental responses elicited by emotional processes conditioned to stimuli accompanying frustrative non-reward. If one effect of septal ablation was to break down a pre-established relationship between S— stimuli and such competing response patterns, then the septal animals should show a selective and immediate increase in post-operative S— responding. Discriminative performance would again emerge as the relationship between S— stimuli, the conditioned emotional processes, and the competing instrumental response patterns was re-established by continued training. Glass, Ison and Thomas (1969) have proposed a similar explanation of the response persistence found after lesions of the anterior limbic cortex in rats. Their hypothesis was formulated to explain why these animals exhibit a strong partial reinforcement extinction effect as well as prolonged responding in extinction after continuously reinforced training.

A consistent finding of both experiments was that septal damage produced a progressive and orderly increase in S+ response rate. The fact that septal lesions do not alter either free-feeding intake (e.g. Donovick, Burright, and Gittleson, 1969) or hunger-motivated free-operant performance when only limited exposure to the schedule is given (Carey, 1969) suggests that the increment is not due to a change in primary drive level. It is also unlikely that the effect is dependent upon the reinstatement of S— responding for the increment is not limited to the context of a discrimination. Dickinson (1972) found a comparable increase in response rate following septal ablations during baseline training when responding was reinforced in both components. However, a cross-study comparison suggests that the increase might be related to the degree of partial reinforcement programmed as well as upon the amount of post-operative training administered. Whilst the present experiments demonstrate a rate increase when a VI-45 s schedule is operative in S+ components, Schwartzbaum et al. (1964) found no change in the S+ response rate of their septal rats when responding was reinforced on a VI-15 s schedule. Also Hothersall, Johnson and Collen (1970) have reported that the magnitude of the septal enhancement of food-reinforced responding on a fixed-ratio schedule increases as the value of the ratio is raised.

In summary, septal damage has been found to disrupt discriminative performance in a free-operant successive discrimination by elevating the S— response rate even when the level of pre-operative S— responding has been minimized by extensive training with either a fixed component-duration or a "correction" procedure. Although the mechanism of the septal action remains unclear, these results suggest that even after prolonged and effective extinction of S— responding, response output is still controlled by an active mechanism capable of disruption by septal damage.

References

AMSEL, A. (1958). The role of frustrative nonreward in noncontinuous reward situations. *Psychological Bulletin*, **55**, 102–19.

AMSEL, A. (1972). Inhibition and mediation in classical and instrumental learning. In BOAKES, R. A., and HALLIDAY, M. S. (Eds), *Inhibition and Learning*, pp. 275–99. London: Academic Press.

AMSEL, A. and WARD, J. S. (1965). Frustration and persistence: Resistance to discrimination following prior experience with the discriminanda. *Psychological Monographs*, **79**, no. 4 (whole no. 595).

BLOOMFIELD, T. M. (1967). Frustration, preference and behavioural contrast. *Quarterly Journal of Experimental Psychology*, **19**, 166–9.

BRADY, J. V. and NAUTA, W. J. H. (1953). Subcortical mechanisms in emotional behaviour: Affective changes following septal forebrain lesions in albino rats. *Journal of Comparative and Physiological Psychology*, **46**, 339–46.

CAPLAN, M. (1970). Effects of withheld reinforcement on timing behaviour of rats in limbic lesions. *Journal of Comparative and Physiological Psychology*, **74**, 467–77.

CAREY, R. J. (1969). Contrasting effects of anterior and posterior septal injury on thirst motivated behaviour. *Physiology and Behavior*, **4**, 759–64.

DALY, H. B. (1971). Evidence for frustration during discrimination learning. *Journal of Experimental Psychology*, **88**, 205–15.

DICKINSON, A. (1972). Septal damage and response output under frustrative nonreward. In BOAKES, R. A., and HALLIDAY, M. S. (Eds), *Inhibition and Learning*, pp. 461–96. London: Academic Press.

DONOVICK, P. J., BURRIGHT, R. G. and GITTLESON, P. L. (1969). Bodyweight and food and water consumption in septal lesioned and operated control rats. *Psychological Reports*, **25**, 303–10.

GLASS, D. H., ISON, J. R. and THOMAS, G. J. (1969). Anterior limbic cortex and partial reinforcement effects on acquisition and extinction of runway response in rats. *Journal of Comparative and Physiological Psychology*, **69**, 17–24.

GRAY, J. A. (1970). Sodium amobarbital, the hippocampal theta rhythm, and the partial reinforcement extinction effect. *Psychological Review*, **77**, 465–80.

HEARST, E. (1971). Contrast and stimulus generalisation following prolonged discrimination training. *Journal of the Experimental Analysis of Behavior*, **15**, 355–63.

HOTHERSALL, D., JOHNSON, D. A. and COLLEN, A. (1970). Fixed-ratio responding following septal lesions in the rat. *Journal of Comparative and Physiological Psychology*, **73**, 470–6.

RILLING, M. E., ASKEW, H. R., AHLSKOG, J. E. and KRAMER, T. J. (1969). Aversive properties of the negative stimulus in a successive discrimination. *Journal of the Experimental Analysis of Behavior*, **12**, 917–32.

SCHWARTZBAUM, J. S., KELLICUT, M. H., SPIETH, T. M. and THOMPSON, J. B. (1964). Effects of septal lesions in rats on response inhibition associated with food reinforced behavior. *Journal of Comparative and Physiological Psychology*, **58**, 217–24.

TERRACE, H. S. (1966a). Stimulus control. In HONIG, W. K. (Ed), *Operant Behavior: Areas of Research and Application*, pp. 271–344. New York: Appleton-Century-Crofts.

TERRACE, H. S. (1966b). Behavioural contrast and the peak shift: effects of extended discrimination training. *Journal of the Experimental Analysis of Behaviour*, **9**, 613–7.

TERRACE, H. S. (1971). Escape from S—. *Learning and Motivation*, **2**, 148–63.

THOMAS, G. J., HOSTETTER, G. and BARKER, D. J. (1968). Behavioral functions of the limbic system. In STELLAR, E. and SPRAGUE, J. M. (Eds), *Progress in Physiological Psychology*, vol. 2, pp. 229–321. New York: Academic Press.

CHARACTERISTICS OF VISUAL AND PROPRIOCEPTIVE RESPONSE TIMES IN THE LEARNING OF A MOTOR SKILL

TIMOTHY C. JORDAN†

Department of Education, Stanford University

After an identical pre-training period 36 subjects were randomly assigned to one of three treatment groups (visual, visual and proprioceptive, proprioceptive, i.e. blindfold) on a motor task. All subjects received three different forces of the stimulus (blade deflection) in a random but balanced fashion. Following 10 training sessions, blindfold subjects had significantly faster RTs. Total response times (RT + MT) were also faster. The three stimulus forces presented produced significant differences in both RT and MT. No differences, however, were found for total response times; a descriptive analysis revealed a step function. As the force increased RT decreased, the opposite effect being observed for MT. Central and peripheral factors were discussed in relation to this phenomenon.

Introduction

Of the various modal inputs to the brain, vision and proprioception, or kinesthesis, are perhaps two of the most important in the area of motor learning.

Evidence from short-term memory codes (Posner and Konick, 1966; Posner, 1967) suggests that visual and kinesthetic mechanisms involve two quite different processing systems, and it has been suggested that persons suffering from perceptual-motor problems (such as apraxia) may find difficulty in the simultaneous integration of these two systems. The flood of incoming information is perhaps too much for them to handle initially. If it could be shown that these systems can be trained independently, then one has a point of departure for working with these individuals.

Also, in a number of highly skilled motor activities responses to a kinesthetic stimulus rather than a visual one would seem beneficial since the kinesthetic route is faster to process (Chernikoff and Taylor, 1952; Botwinick and Brindley, 1962; Keele and Posner, 1968; Gibbon and Rutshmann, 1969).

Kinesthetic cues, however, are seldom the only means through which the subject perceives his movements, indeed for the beginner there is considerable evidence that when both visual and kinesthetic stimuli occur together, the visual tends to be dominant in its demands for attention (Gibbs, 1954). One is inclined to rely upon what one sees, rather than upon what one feels. Thus in skills

†Present address: Department of Experimental Psychology, University of Oxford.

where kinesthetic or proprioceptive† cues are considered important, this can perhaps be a hinderance to the learning of an optimal response.

Although learning has been studied as a function of the type of sensory stimuli available, there are still many unanswered questions as regards proprioception, particularly concerning the kinds of perception and adjustments occurring at relatively unconscious levels. Pew (1970) notes that "we talk of the importance of proprioceptive feedback but it is not clear how to include this class of information into a representation of skilled performance . . . we need to delineate the processing stages involved and to examine their roles in the conduct of a skilled task".

The purpose of this investigation was to examine the phenomena involved in the acquisition of a motor skill when visual feedback was withheld. It was intended to demonstrate that by such deprivation the subject would more easily focus his attention on the proprioceptive feedback, thus producing faster responses. The relationships between reaction time (RT) and movement time (MT) with respect to these two modalities were also studied in an attempt to clarify the somewhat inconclusive evidence in this field.

Method

Subjects

Subjects for this investigation were 36 Stanford University students (24 female and 12 male, with a mean age of 20·61 years). The majority of subjects received partial pscyhology credit for their participation. The remainder were volunteers and paid volunteers. All subjects were naive fencers.

Apparatus and instrumentation

Equipment used for this study consisted of (i) surface electrodes with telemetry connectors for EMG recording; (ii) physiograph PMP-4A with biotelemetry receivers type FM-1100-6 using cardiac preamplifiers; (iii) electrical epees; (iv) blindfold masks; (v) variable force blade deflection mechanism (stimulus).

The task was a fencing skill, the disengage and lunge. The stimulus used was the deflection of a mechanical blade (substituting for the parry of an opponent). On presentation of the stimulus, subjects were required to disengage their blade by a controlled finger and wrist movement bringing their blade under and around the mechanical foil, extending their arm and hitting a target pad mounted on the facing wall immediately beyond the stimulus mechanism [see Figs 1(a) and (b)].

To standardize the speed and force of the deflection stimulus, a special piece of equipment was constructed. This consisted of a foil blade mounted on a pivot [see Fig. 1(a)]. The deflection was created by using a solenoid and control was gained by the use of a spring balance. The foil's handle having been removed, the solenoid and spring balance were coupled to opposite sides of the tang. The experimenter pressed the control button at the beginning of each trial. After a random time delay of between ½ and 3 s at each trial, the current drew the bolt into the solenoid deflecting the blade against the resistance of the spring

---

†In this study the term proprioceptive is preferred over kinesthetic, since strictly speaking kinesthetic refers to the conscious sense by which limb position and motion are perceived; i.e. information received for the most part from the joint capsules. As a result of the stimulus presentation used in this study, one cannot be certain that only these receptors are responsible for the responses observed.

balance. The solenoid strength being known, the spring balance was set at three separate positions thus providing three different initial forces of blade deflection—$\frac{1}{2}$ lb, 1 lb, and 1$\frac{1}{2}$ lb (0·226 kg, 0·453 kg, and 0·680 kg) respectively.

RT was defined as being the period between the onset of the stimulus and the change of action potential in the voluntary muscles. Recordings were made from the flexor group of muscles of the fingers, controlling the disengage, on a line from the medial epicondyle of the humerous to the head of the ulna. Tests were made to determine whether deflection of the blade alone (with no response) caused observable firing to be recorded from the muscles. No such interference was found, thus a clear separation of RT and MT was readily available.

MT was therefore the time taken to perform the complete movement of the disengage and lunge measured from the first burst of firing in the voluntary muscles until the target pad was struck.
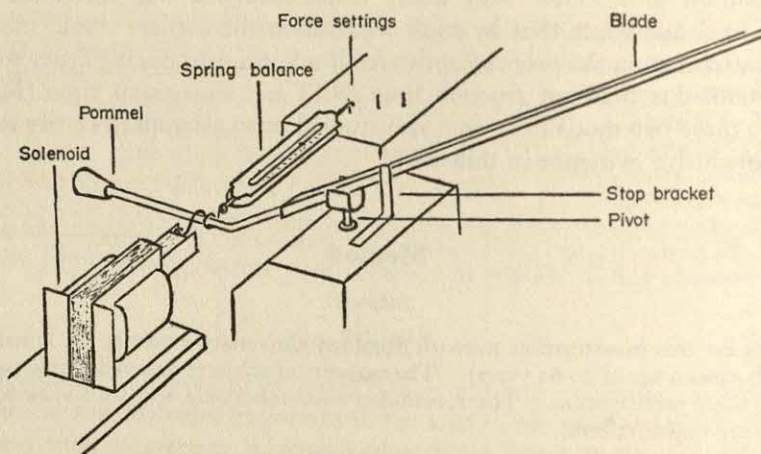
FIGURE 1(a). Blade deflection mechanism.
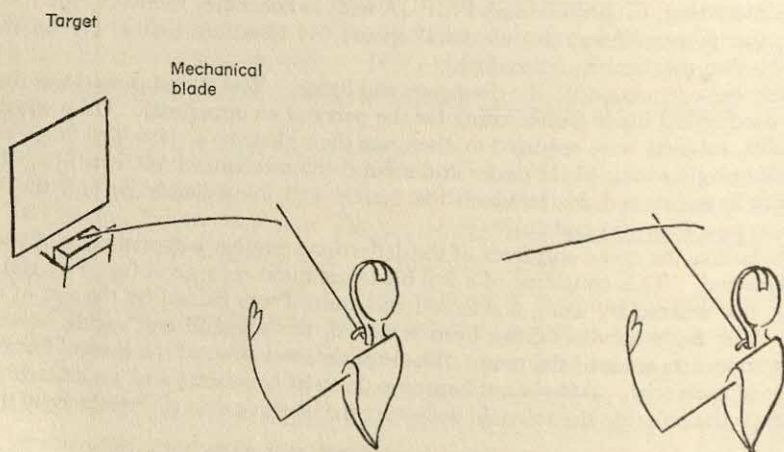
(i) "On Guard" - no contact   (ii) − with blade contact

(b) Target and starting positions.

The target consisted of a 2 ft (61 cm) square padded area. When the pad was struck by the subject's blade, an electrical circuit passing via the blade was completed and a pulse was recorded on the physiograph thus marking the termination of MT.

Recording of all measures, time and event and EMG, was done on the physiograph with the paper control set at 5 cm/s.

Procedures and design

Following a pre-training period, in which the fundamentals leading to the fencing lunge with a disengage were practised, the full lunging distance (fencing measure) of each subject was determined. This ensured a standardization of starting position and distance from the target. Mean fencing measures for each group were 2·447, 2·468 and 2·478 m respectively.

Subjects (with equal numbers of male and female) were randomly assigned to one of the three treatment groups and were tested over a period of 10 training sessions under one of these conditions. Subjects in group I came "on guard", sighted, with their blades approximately 6 in. (15 cm) away from the mechanical foil. That is, their initial information as regards the stimulus was entirely visual. Subjects in group II could also see and came on guard with their blades in contact with the mechanical foil, thus having both visual and proprioceptive information. The third group had their blades in contact with the mechanical foil but were blindfolded [see Fig. 1(b) for relative starting positions].

Each group received 15 trials during a test session in order to allow for a balance of the three forces of deflection presented.

Subjects were instructed to respond as *quickly* and *efficiently* as possible to the deflection of the mechanical blade (stimulus).

Results

Treatments

Analysis of variance on the total response times obtained in test 10 indicated that the blindfold group were significantly faster ($P < 0·05$) in their responses. The variance for this group was also much smaller than for either of the sighted
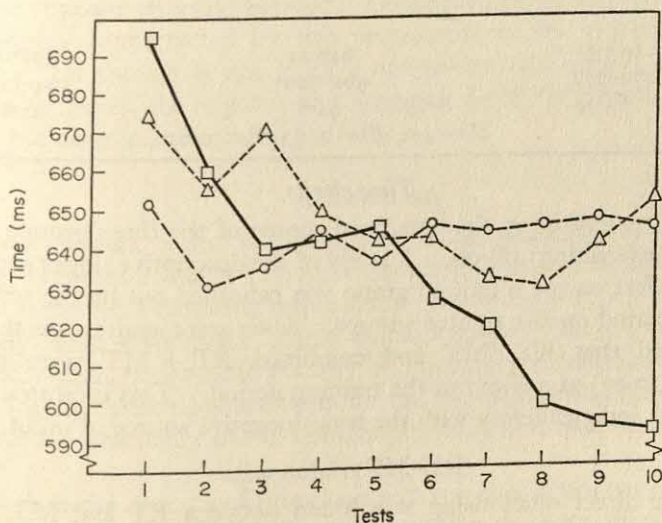
FIGURE 2. Group means for Combined RT + MT on Tests 1–10. △ Group I; ○ Group II; □ Group III.

groups. No significant differences in response times were found between the sighted groups.

It would seem, therefore, that withholding visual input in the early learning stages of a motor skill allows the subject to focus his attention on the remaining proprioceptive source of input, thus producing faster total responses.

Reaction times for the blindfold group were significantly faster ($P < 0.01$), and there was no overlap in the range of scores between the sighted and blindfold group. This would not only seem to confirm the dominance of the visual input when present, but also indicate that the blindfold subjects are working on quite a different system. Similar significant differences ($P < 0.01$) were observed for each of the three separate forces presented.

Movement times, though markedly reduced, were not significantly different ($P > 0.01$) owing to the high variance observed in this phase of the task (see Table I for a summary of mean response times on test 10).

<div align="center">TABLE I</div>

<div align="center">*Summary of response times on test 10 (ms)*</div>

	Group I (S)	Group II (S + P)	Group III (P)
RT			
X̄	129·16	136·01	109·11
Range	120–143	121–174	96–118
S.D.	6·71	15·56	6·42
		$F = 21·48, df = 2, 33, P < 0.01$	
MT			
X̄	524·33	510·92	483·98
Range	348–646	433–580	450–524
S.D.	82·85	40·41	24·04
		$F = 1·69, df = 2, 33, P > 0.05$	
RT + MT			
X̄	653·5	647·21	592·09
Range	474–777	561–709	564–629
S.D.	85·77	44·8	24·07
		$F = 4·0, df = 2, 33, P < 0.05$	

Time effects

Analysis of regression on the combined means of the three groups showed no significant time (training) effects. A study of the descriptive slopes suggested that any marked effect on the blindfold group was cancelled out by the weak effect of the training period on the sighted groups. A separate analysis on the blindfold group revealed that RT, MT and combined RT + MT were significantly reduced ($P < 0.01$) as a result of the training period. This indicated an obvious ability to work independently with the proprioceptive source of input.

RT–MT relationships

A very low direct relationship was found between RT and MT ($r = 0.28$); however, a more detailed analysis of the effects of the force variables did reveal a different picture.

Force variables

Analysis of variance on test 10 to determine the effects of the different forces presented revealed significant effects ($P < 0.01$) for each of the dependent variables RT and MT. For the combined RT + MT response no significant difference was observed at this level.

TABLE II

Summary of group means for force variables in test 10

Forces		1	2	3	
RT	Group I	134·50	127·83	125·16	
(ms)	II	142·41	136·45	129·16	
	III	113·16	111·91	102·25	
	Mean	130·02	125·01	118·86	$F = 37.4, df = 2, 66, P < 0.01$
MT	Group I	516·66	518·16	538·16	
(ms)	II	505·39	506·54	521·66	
	III	472·95	482·58	496·41	
	Mean	498·34	502·43	518·75	$F = 14.39, df = 2, 66, P < 0.01$
RT + MT	Group I	651·16	646·00	663·33	
(ms)	II	647·81	643·00	650·83	
	III	586·12	594·50	598·66	
	Mean	628·36	627·83	637·61	$F = 3.74, df = 2, 66, P < 0.05$

A descriptive analysis of these findings (summarized in Table II) provides clarification regarding these results. It can be seen that for each difference in force, a corresponding response is created. For RT, as the force increases the response times decrease; that is, the stronger force produces the faster response. However, for MT the opposite effect is observed. Consequently, the force effect at any given time is largely counteracted by this phenomena in the combined RT + MT response. This though is not quite a one-to-one relationship, since the mean differences (between the weakest and strongest force deflections) are longer for MT than the corresponding ones for RT.

Discussion

The basic findings, within the limitations of this study, would seem to support the initial hypothesis that a period of blindfold practice in the early learning stages of a motor skill requiring responses to proprioceptive feedback will produce significantly faster responses. These findings support Keele (1968) in that for predictable events, movement control is apparently internalized in the CNS. The significantly faster response times for the blindfold group provides additional evidence of the dominance of the visual input in its demands for attention by the brain.

Perhaps then the concept of proprioceptive response times needs readjusting in terms of ultimate values. This is not to say that previous measures have been inaccurate, merely that this study suggests that in the normal person responses to

proprioceptive stimuli seem less than optimally developed owing to the tendency of domination by other sources of input. That is unless one considers them over a period of time, where an attempt is made to focus on a maximal response one is not necessarily obtaining the truest measure.

From the summary of data in Table I one can see that mean RTs were very fast, certainly faster than might have been anticipated from previous observations (Chernikoff and Taylor, 1952; Keele and Posner, 1968). Chernikoff and Taylor measured the time taken for subjects to arrest the dropping of their arm from the horizontal position. They used an accelerometer to indicate changes in acceleration of the arm once it was released. This change in acceleration was taken as the measure of kinesthetic RT. In other studies finger lift responses or discrete tracking movements have been used. This investigation appears to be the first time that applied variable forces have been used as stimuli for RT. It is conceivable that this factor alone could account for some of the differences, especially since it was shown that different degrees of force result in different speeds of RT.

The main discrepancy in these values, however, probably lies within the definition of RT as used in this study. Recent researchers (Weiss, 1965; Botwinick and Thompson, 1966) have divided RT into two components, namely premotor and motor RT. Premotor RT (pmRT) was defined as being the time from the onset of the stimulus to the change in EMG, and the motor component (mRT) the time from EMG change to the finger lift response. The pmRT correlated very highly with total RT and showed comparable variations as a function of differing preparatory intervals. It was this same interval (from stimulus until EMG change) which was defined as RT for this study. mRT was poorly correlated with total RT and appeared to be independent of the preparatory interval.

Clarke (1968) found a reduction in total RT as the level of pretension was increased on a hand grip device. The present findings indicate a similarity with Clarke's findings to the extent that the stronger the force exerted, the faster were the RTs recorded. In attempting to explain his phenomenon Clarke considered the possibility of changes in the local musculature. He also considered the possibility of the locus being central rather than peripheral: subjects could be increased in their arousal state as a result of increased proprioceptive feedback. This would account for the three distinctive speeds of RT in the present investigations and would also suggest a close parallel with those of Lissman (1950) who noted an increased discharge in some primary endings in isometric conditions. It does not, however, account for the opposite increase in the corresponding MTs.

An initial explanation considered was that as the force of blade deflection increased (causing a decrease in pmRT) the subject's blade was deflected further out of line, necessitating a proportionately greater route to the target, hence the corresponding MT would be increased. This, however, is not an entirely satisfactory explanation for group I (visual only) who had no initial contact with the blade. They were in fact rarely ever able to "beat the blade" and avoid any contact, consequently some degree of proprioceptive information was potentially available. As the time for complete (mechanical) blade deflection was in the order of 70–90 ms, it seems most unlikely that they were waiting for this additional source of information before responding.

Schmidt and Stull (1970) in seeking the locus of Clarke's change in total RT found no significant difference in total times, but did find that as pmRT decreased the mRT component increased. Partial programming or activation hypotheses were tenable for the pmRT changes, whereas changes in the contractile component were tenable for the mRT changes.

Whether the differences between pmRT and mRT observed by Schmidt and Stull can completely account for the effects observed in this investigation are difficult to determine. They found little or no change in total RT. A descriptive analysis of the respective differences between RT and MT in this study suggests that the decreases in RT steps are more than compensated for in the step differences between MT. This could be a cumulative effect of Schmidt and Stull's observations, together with some slight increase in blade deflection. Second, MT in this instance includes both the mRT component and the ouvert movement response.

Further study is clearly indicated, with a division being made between pmRT, mRT and MT, before any definitive theoretical explanation or model describing these findings can be forwarded.

References

BOTWINICK, J. B. and BRINDLEY, J. F. (1962). An analysis of set in relation to reaction time. *Journal of Experimental Psychology*, **63**, 568–74.

BOTWINICK, J. B. and THOMPSON, L. W. (1966). Premotor and motor components of reaction time. *Journal of Experimental Psychology*, **71**, 9–15.

CHERNIKOFF, R. and TAYLOR, F. V. (1952). Reaction time to kinesthetic stimulation resulting from sudden arm displacement. *Journal of Experimental Psychology*, **43**, 1–8.

CLARKE, D. H. (1968). Effect of preliminary muscular tension on reaction latency. *Research Quarterly*, **39**, 60–6.

GIBBON, J. and RUTSHMANN, R. (1969). Temporal order judgement and reaction time. *Science*, **165**, 413–5.

GIBBS, C. B. (1954). The continuous regulation of skilled response by kinaesthetic feedback. *British Journal of Psychology*, **45**, 24–9.

KEELE, S. W. (1968). Movement control in skilled motor performance. *Psychological Bulletin*, **70**, 387–403.

KEELE, S. W. and POSNER, M. I. (1968). Processing of visual feedback in rapid movements. *Journal of Experimental Psychology*, **77**, 155–8.

LISSMAN, H. W. (1950). Proprioception. In *Physiological Mechanisms of Animal Behaviour*. Experimental Biology Symposium no. IV. Cambridge University Press.

PEW, R. W. (1970). Toward a process-oriented theory of human skill performance. *Journal of Motor Behaviour*, **2**, 8–24.

POSNER, M. I. (1967). Characteristics of visual and kinesthetic memory codes. *Journal of Experimental Psychology*, **75**, 103–7.

POSNER, M. I. and KONICK, A. F. (1966). Short-term retention of visual and kinesthetic information. *Journal of Organizational Behaviour and Human Performance*, **1**, 71–88.

SCHMIDT, R. A. and STULL, B. (1970). Premotor and motor reaction time as a function of preliminary muscular tension. *Journal of Motor Behaviour*, **11**, 96–110.

WEISS, A. D. (1965). The locus of reaction time change with set, motivation and age. *Journal of Gerontology*, **20**, 60–4.

EFFECTS OF ELECTRIC SHOCK ON EXPLORATORY BEHAVIOUR IN THE RAT

D. I. WILLIAMS

Department of Psychology, The University, Hull

Rats were given five daily 3-min trials in a Y-maze with two grey and one black-and-white striped arms. An initial preference was shown for the striped arm. Between Trials 3 and 4 half the subjects were given electric shock, half in a striped box and half in a box lined with aluminium foil. Shock resulted in reduced ambulation in the maze, but increased the preference for the striped arm in terms of initial choice and time spent in it. There was no significant difference between the two shock conditions. Results are consistent with the notion that mild fear may facilitate exploration.

Introduction

Halliday (1966) suggested that rats may explore stimuli because they evoke fear rather than because they are novel. This notion has been tested by comparing performance in elevated and enclosed mazes (Halliday, 1967) which are supposed to evoke different amounts of fear, but it is possible that these situations differ in novelty (Sheldon, 1968). An unambiguous test of Halliday's position requires that fear levels be manipulated by some means which does not affect novelty. This has been done (Halliday, 1966) in a study which showed that exploration of the striped section of a maze was increased by shocking rats in a box with striped walls. This is an important result as it is in direct contrast with the general expectation that exploration is depressed by aversive stimulation (Baron, 1963; Kumar, 1968). Kumar (1970) failed to replicate Halliday's result; he showed that animals avoided stimuli associated with shock. This does not invalidate Halliday's position as it is clear that at certain shock levels avoidance would be predicted. Only at mild fear levels will exploration be facilitated. The present study is a further attempt to replicate Halliday's (1966) finding.

Method

Subjects

Forty-four male black-hooded rats (strain PVG/C) aged approximately 100 days were used.

Apparatus

A symmetrical enclosed Y-maze was used, with arms 46 cm long, 10·5 cm wide, and walls 22 cm high. It was painted in flat grey except for one arm which had regular black and white vertical stripes, 2·5 cm wide. A thin black line was drawn across the entrance to the

two grey arms. The maze was placed in a rigid enclosure painted flat grey and measuring 141·5 × 141·5 × 123·5 cm high; one side was open to allow entry and observation. The maze was illuminated by a single 60-W bulb suspended centrally from the top of the enclosure. In a separate room to the maze was a shock box which measured 22·5 × 22·5 × 19·0 cm with a grid floor; the walls were either covered in aluminium foil or painted in regular black and white vertical stripes 2·5 cm wide. The shock box was illuminated from above to give approximately the same illumination as found in the maze.

Procedure

Rats were given five 3-min trials in the maze with an inter-trial interval of approximately 24 h. Measures were taken of defecation (number of boli deposited), number of arms entered and the order in which they were entered. A record was also made of the time spent in the striped arm. An arm entry was scored when the head and three of the animal's feet had crossed the line at the entrance to the arm. 2 h after the end of Trial 3 each rat was placed in the shock box for 30 s. For half the animals the interior was striped, for the rest the lining was of aluminium foil. Half the animals in each condition were unshocked, and half shocked, receiving three scrambled shocks (1·6 mA) of 0·5 s duration delivered through the feet at 5-s intervals during the last 20 s in the shock box.

Results

An analysis of variance on arm entry scores by the four groups on Trials 1–3 (prior to shock) showed no significant differences except for Trials ($F = 44·42$, $df = 2, 80$, $P < 0·001$) indicating a drop in the number of arms entered over trials. A similar analysis on Trials 4 and 5 (after shock) showed significantly fewer arm entries by the shocked groups ($F = 8·73$, $df = 1, 40$, $P < 0·01$), but with no difference between the striped and foil conditions ($F < 1$), and no significant interactions ($F < 1$); the effect of Trials was significant ($F = 11·45$, $df = 1, 40$, $P < 0·01$) indicating an increase in arm entries in all groups from Trial 4 to Trial 5. The total number of arm entries for striped and foil conditions respectively were for Trial 4, shock: 25, 24; no shock: 41, 37, and for Trial 5, shock: 41, 37; no shock: 58, 47.

An analysis of variance of time spent in the striped arm prior to shock (Trials 1–3) showed no significant differences between the groups, although on each day animals spent longer than chance expectation in the striped arm ($z = 3·17$, $P < 0·001$). Over the 3 trial period an average of 63% of the time was spent in the striped arm. After shock (Trials 4 and 5) the shocked groups spent significantly more time in the striped arm ($F = 8·18$, $df = 1, 40$, $P < 0·01$); no other effects were significant. On Trial 4 shock groups spent 84% (striped) and 91% (foil) of the time in the striped arm compared with 52% (striped) and 59% (foil) for the unshocked controls. Both shock groups stayed longer than the chance expectation in the striped arm on Trials 4 and 5 ($P < 0·001$, binomial test), whereas this was not so in either unshocked control ($P > 0·05$, binomial test). There were no significant differences in the proportion of entries into the striped arm on any trial, although owing to the small number of arm entries by the shock groups after shock this measure is not very meaningful.

An analysis of variance of defecation scores on Trials 1–3 showed no significant difference between the groups. On Trials 4–5 after shock, the shock groups were seen to defecate significantly more than the unshocked groups ($F = 4·08$, $df = 1, 40$, $P < 0·05$).

Discussion

From the increased defecation scores for the shock groups on Trials 4 and 5 it may be inferred that shock had the desired effect of stressing the animals. The result was to reduce the amount of movement in the maze, but to increase the preference for the striped arm as measured both in terms of initial choice of that arm and the time spent in it. The preference measure provides an unequivocal measure of what is being explored; thus it seems that in general Halliday was correct in asserting that exploration may be facilitated by fear. Rats, on initial contact with the maze spent more time in the striped arm, and chose to enter it on leaving the start arm. This behaviour is reinstated and enhanced following shock.

The foil condition was intended to provide an environment as different as possible from both the striped and grey arms of the maze. As exploration of the striped arm was facilitated by the animals being shocked in the foil box, it seems that fear has a general facilitatory effect. Halliday's specific finding that shock in the presence of a particular stimulus enhances exploration of that specific stimulus was not then replicated.

References

BARON, A. (1963). Differential effects of fear on activity in novel and familiar environments. *Psychological Reports*, **13**, 251–7.

HALLIDAY, M. S. (1966). Exploration and fear in the rat. *Symposium of the Zoological Society of London*, **18**, 45–59.

HALLIDAY, M. S. (1967). Exploratory behaviour in elevated and enclosed mazes. *Quarterly Journal of Experimental Psychology*, **19**, 254–63.

KUMAR, R. (1968). Psycho-active drugs, exploratory activity and fear. *Nature*, **218**, 665–7.

KUMAR, R. (1970). Effects of fear on exploratory behaviour in rats. *Quarterly Journal of Experimental Psychology*, **22**, 205–14.

SHELDON, M. H. (1968). Exploratory behaviour: the inadequacy of activity measures. *Psychonomic Science*, **11**, 38.

BOOK REVIEWS

PIAGET, J. *Structuralism.* London: Routledge & Kegan Paul. 1971. Pp. 153. £2.00.

This is an amusing book. It is, actually, the first do-it-yourself apotheosis that I have ever come across in Psychology. It is about Structuralism, and here Piaget's definition of a structuralist seems to be anyone who is not a S-R psychologist, anyone, that is, who is concerned with any kind of organization. Piaget considers Chomsky. Chomsky is judged as a very good structuralist apart from the fact that he is not sufficiently concerned with the development of structures. The Gestalt psychologists are also praised for their recognition of the importance of structures, but they too fail to treat development seriously enough. Lévi-Strauss is shown to have made his contribution, but he, alas, is again too static. The last chapter makes everything clear. There are two properly developmental structuralists. One is Piaget; the other is God. "But God himself, since Godel's theorem, has ceased to be motionless. He is the living God, more so than heretofore, because he is ceaselessly constructing ever stronger systems" (p. 141). I wonder now who it is that we have been waiting for. Is it for God or for Gödel or for Piaget?

<div align="right">P. E. BRYANT</div>

ELKIND, D. *Children and Adolescents: Interpreting Essays on Jean Piaget.* London: Oxford University Press. 1971. Pp. 160. £2.70.

Nowadays even the toughest experimental psychologist knows about Piaget's work on cognitive development, and this is fine because experimental psychologists, on the whole, can understand what is wrong as well as what is right about a series of experiments. They will know, or will easily be persuaded that, while the questions which Piaget asks about development are very stimulating, his approach to experimental design is distinctly cavalier (in the "1066 and All That" sense of being wrong but romantic) and that his actual experiments are rather chaotic.

The problem really is that Piaget's work is known about by other people as well. In fact he is probably now the best known psychologist since Freud. It is a problem because his work is accepted quite uncritically by people whose interests are practical and who tend to be rather irritated by questions of experimental controls. In fact the most worrying thing about Piaget is the wide and unquestioning acceptance which his work has had in colleges of education and the quite unjustified educational conclusions that have been drawn from it there.

So it is depressing to see yet another uncritical book on Piaget's work. David Elkind has put together a series of nine "interpretive essays" on Piaget. They make a very low-level book, and I am surprised that any university press (let alone Oxford's) should publish this sort of production. The first essay is like something out of a fan magazine. Piaget is nice to children, has a mane of white hair, smokes a meerschaum pipe and rides a bike (don't we all?). From adulation we go to exposition in the next eight chapters. One is called "How the mind grows" and in it we are told, among other things, what the child answers when he is asked if a dog or a cat could be a Jew or a Protestant. I wonder whether this sort of question is really necessary. The remaining chapters are mainly concerned with Piaget's relation to education and to aptitude tests. These are sensible essays but they do not tackle the basic question of whether Piaget's theories about development are justified or not. Surely this is the important question, which should be considered by educationalists as much as by psychologists. Yet this book implies that the question does not exist.

<div align="right">P. E. BRYANT</div>

OLSON, D. R. *Cognitive Development: the Child's Acquisition of Diagonality*. New York: Academic Press. 1970. £5.60.

It is very clear that young children, as well as octopuses, rats and some fishes, find it very difficult to distinguish one oblique line from another. It is really quite intriguing that this difficulty should occur so strongly with such a wide range of species, especially since it applies very specifically to discriminations between obliques and not to discriminations between, for example, horizontal and vertical lines.

Any explanation of the specific difficulty which young children have with obliques must deal with two distinctions. The first is between simultaneous and successive displays. The difficulty which young children have with obliques occurs only in successive tasks. Children make scarcely an error when their job is to match simultaneously presented oblique lines. The second distinction is between mirror image and non-mirror image configurations. It is often said that obliques are difficult because they are mirror images, that is to say because they are symmetrical around the vertical axis, and that horizontal-vertical pairs are easy because they are not symmetrical in this way. The obvious test of this hypothesis is to present discriminations between obliques which are asymmetrical around the vertical axis, and yet strangely experiments with young children have avoided this test.

Olson's book is entirely concerned with children's difficulties with obliques, and yet it manages to ignore both these crucial issues. It is an account of a series of experiments (five years' work) in which children were given problems which involved either constructing or matching obliques. Sometimes the problems were simultaneous, sometimes successive, and they always involved just mirror image obliques. The main conclusion that can be drawn from these experiments, as far as I can see, is again that obliques are peculiarly difficult for young children.

Olson's own hypothesis is not very specific. He assumes, unjustifiably, that the difficulty is a mirror image one. His theory about how the children eventually overcame the difficulty is not really based on any hard data and is couched in the most terrible jargon. They learn to distinguish between obliques by making appropriate "performatory acts". This horrible phrase crops up at least three times a page throughout the long concluding chapter. I hope that no other psychologist ever takes it up.

P. E. BRYANT

MENYUK, P. *Sentences Children Use*. Cambridge, Mass.: MIT Press. 1969. Pp. xiv + 165. £3.50.

Despite a ritual bow to theoretical linguistics, *Sentences Children Use* lies very much within the descriptive tradition of language development studies. Granted the basic assumption that children are acquiring the rules of English grammar, Menyuk gives a straightforward account of possible rule structures underlying the utterances of children ranging from two to seven.

It is instructive to compare this research with other more fashionable reports of child grammars, notably those by Roger Brown, David McNeill and their associates. On the credit side, Menyuk's lack of theoretical dogma and tentative analysis of the data leave one with the impression that, for once, the children's utterances have been allowed to speak for themselves. On the other hand, her "common-sense" approach ignores many crucial distinctions relevant to current linguistic controversies. To take just one example, use of adjectives is at one point treated as resulting from expansion of base structures and at another as requiring mastery of a generalized transformation, thus befogging the whole issue of whether there is a pre-transformational stage in child speech. Similarly, no clear basis is given for distinguishing between the development of syntactic and semantic selectional restrictions on word co-occurrences.

Due to this lack of theoretical bite, the opportunity is lost of testing various linguistic theories about deep and surface structure, syntactic and semantic rules and innate linguistic universals against the recorded utterances of an unusually large sample of children. The exception is an interesting chapter in which procedures for eliciting linguistic competence

are used to test hypothesized differences between normal and deviant speaking children's ability to repeat sentences and lists of unrelated words. The overall impression, however, is of undigested empirical data, presented with an unfortunate monotony of style which perhaps accounts for the relative lack of interest in this monograph since it first appeared three years ago.

<div style="text-align: right">JUDITH GREENE</div>

SLOBIN, D. I. (Ed.). *The Ontogenesis of Grammar*. New York, London: Academic Press. 1971. Pp. 247. £5.85.

HUXLEY, R. AND INGRAM, E. *Language Acquisition: Models and Methods*. New York, London: Academic Press. 1971. Pp. 311. £5.00.

BLUMENTHAL, A. L. *Language and Psychology: Historical Aspects of Psycholinguistics*. New York: John Wiley & Sons, Inc. 1970. Pp. 248. £3.75.

The Ontogenesis of Grammar is an attempt by theorists of several different camps to offer an interpretation of the same body of data. The data, presented by Slobin, consist of familiar observations of children's speech: the early two- and three-word utterances, overgeneralization of regular tense forms in English and various stages in the development of the adult form of negation.

McNeill, in a short chapter, makes immediate reference to "innate linguistic capacities". Given the task he has set for the child this interpretation is inescapable: for McNeill the infant is a syntactic analyser, a miniature grammarian whose activities act as a metaphor for the processes of acquisition. The difficulty is, of course, that the task real children perform may be something quite different, involving a multitude of non-linguistic acts and perceptions, and building on a wealth of earlier conceptual development. The conclusion, that grammatical classes arise in the derivation of syntactic features from basic grammatical relations, when stripped of its nativist foundations leaves all the interesting questions unanswered. Palermo attempts to show the relevance of principles developed in the traditional areas of verbal learning to language acquisition, but in this he is singularly ineffective, making do with a rather half-hearted appeal to associative chaining and the progressive refinement of utterances through reinforcement, with a goal of mutual intelligibility. Schlesinger argues to replace the traditional P-markers of transformational grammar with I-markers (input-markers) reflecting the underlying intentions of the speaker. The child learns a grammar by learning how to place words relative to each other so as to express certain relations: in production, meaning comes first and the child learns to realize this in linguistic form. Comprehension is the process of production reversed.

The chapter by Staats is the longest in the book and could have been severely pruned. He begins with a plea for a better understanding of the vast range and variety of learning theories, which makes any sweeping condemnation of "learning theory" in general absurd, but then commits just this crime in a bitter and uncontrolled attack on a group of theorists, styled "Chomskian", whose defining characteristics appear to be an entrenched nativism and proclivity for circular argument. Precisely who the adherents of this group are is left to the reader's imagination. The argument directed at Chomsky as a linguist takes the form that information about a language gained from observation of native speakers cannot "explain" their behaviour, which is neither a novel comment nor, in view of the competence/performance distinction, particularly well aimed. The constructive contribution in the chapter is an elaboration of the author's well-known variety of mediational theory, but where this is applied to the data set out by Slobin it is unconvincing.

The book ends with two post-symposium contributions. The first, by Braine, offers a brief and elegant examination of Chomsky's notion that the child tests hypotheses "to determine which of the (humanly) possible languages is that of the community in which he is placed". The second contribution, by Erwin-Tripp, exactly summarizes what is wrong with the book as a whole: she remarks that time has "left some authors guarding outdated positions", and indeed late revisions of these papers, initially written for a 1965 symposium,

have done little or nothing to revitalize them. It is difficult to recommend a book containing so many outmoded views presented by and large in a style of unyielding polemic.

Language Acquisition: Models and Methods could so easily have been a similar book, representing the collected papers of a Study Group on Language Development held as long ago as 1968, but it is not. Huxley and Ingram are to be congratulated on bringing together an excellent representation of contemporary work in psycholinguistics. The book falls into four sections: Sociolinguistics, Grammar, Cognition and Language, and Clinical. Representing Sociolinguistics, Hymes, Erwin-Tripp, Cazden, Robinson and Bernstein all press the case for a consideration of language usage, "competence for use", as distinct from the preoccupation of some linguists with the ideal speaker/hearer. As Hymes puts it: we need, and do not have, a generative theory within which socio-cultural factors have an explicit role. The real children of New York are a far cry from the unfolding fluent child of linguistic theory. Erwin-Tripp extends this argument by a detailed examination of non-standard negro English and introduces the notion of "socio-linguistic rules" allowing for the influence of factors such as styles of address and the acceptability (as distinct from grammaticality) of language in different contexts. Cazden, Robinson and Bernstein conclude the section with appeals for a widening of the scope of linguistic enquiry to include semantic considerations (broadly defined), and the intentions of speakers.

The second section of the book is dominated by a closely argued paper by Sinclair attempting to show the importance of sensorimotor intelligence as a precondition for the acquisition of syntax. The pre-verbal child, in this view, possesses equivalent operations to the concatenation, classification and categorization of linguistic theory, together with cognitive equivalents of such grammatical relations as actor-action-object. In general she argues her case well but the reader is left with certain residual doubts as to the exact mechanisms which permit the non-linguistic conceptual schemas to find an eventual linguistic form. The contribution by Ingram points out the problems posed for psychologists by the frequent changes in linguistic theory: the psychological implausibility of certain assumptions (e.g. the biogenetic preprogramming of LAD); its isolation from "real" language; and the general incompatibility of some of the concepts with current psychological theory. The issue is perfectly encapsulated in Bever's contention that in one sense linguistic intuition cannot be refuted experimentally, since "if an empirical test shows you what you know is true, is false, then there is something wrong with the empirical test".

In the section on Cognition and Language, among contributions by Church and Mehler a long paper by Bever illustrates the development of cerebral dominance in one of the "perceptual strategies" he has recently elaborated elsewhere. The evidence shows that the NVN strategy (treating all noun-verb-noun sequences as actor-action-object descriptions) is not immediately available to the non-dominant ear-hemisphere. The conclusion, "that cerebral lateralization is itself critically dependent on certain kinds of experience", will no doubt generate further socio-cultural studies. As an intellectual bonus for the reader the book presents an edited transcript of the discussion following each paper, a device which must rarely have succeeded so well.

Language and Psychology is a short and highly selective treatment of some of the historical background to modern psycholinguistics, inspired, in the author's words, by the fact that "American psychologists are seemingly unaware of an illustrious earlier history". The book must be judged as only partly successful. Blumenthal's choice of extracts seems curiously arbitrary, and there are many notable omissions. The extracts from recent writers (Chomsky, Miller, Lashley and Lenneberg) are all very familiar and easily available elsewhere. It is difficult to see for whom this book was written, since the serious historian will find it eccentric and unbalanced, and as an introduction to the subject it is unnecessarily obscure. A history of psycholinguistic research must surely be written—but this is not the book.

R. A. KENNEDY

BERNOTAT, R. K. AND GARTNER, K.-P. (Eds). *Displays and Controls: Proceedings of a NATO Advanced Study Institute held at Berchtesgaden, Germany, in 1971.* Amsterdam: Swets & Zeitlinger. 1972. Pp. 493. $28.

Given the title, it is ironic that this book is a prime example of how *not* to display information. The collected papers seem to be photocopies of the originals; appropriate figures are tucked in at the end of each paper rather than being suitably embedded in the text; and some of the German-English translations resemble zero-order approximations to the latter language. In particular, one feels that more attention could have been paid to the included paper by Singleton on information presentation.

Those who bring the psychologist's approach to problems in human engineering are unlikely to derive anything of value from the collection. For them, the title is sadly misleading since the main theme is not about optimization in the design of displays and controls. Instead the bias is towards the use of pre-displays to improve tracking performance and to the development of further descriptive, control-theoretic models of the human operator. These models, given structure in analogue and/or digital devices, admirably simulate human tracking performance. Since they will operate in a "faster than real-time" mode, they can provide prediction displays to guide the human operator and so improve his performance. Strangely enough, it is only at the end of a paper by Johannsen on his non-linear multi-parameter model that we get even a glimpse of the obvious implications.

The more psychologist-oriented papers are a disappointment because they contain nothing that is new: Haider, for example, reminds us of that 1964 paper; Singleton points out that experiments on information presentation are really quite varied (which we already knew) and that a psychophysiological approach might prove useful (which some of us already knew, too); and Vicki Cohen, in a classic "old wine in new bottles" paper, describes work on digital, moving pointer and moving scale displays.

The real value of this volume may well lie in the example it sets for future editors on how not to edit conference proceedings.

A. CRAIG

YOUNG, J. Z. *An Introduction to the Study of Man.* Oxford: Clarendon Press. 1971. Pp. 719 + xxv. £6.00.

The author tells us in the preface that this book arose out of a series of lectures given to medical and dental students at the beginning of their studies. They were fortunate students indeed, and one must be grateful to the author for expanding his notes into this excellent volume for the benefit of a wider audience.

The range of topics is immense: from the chemistry of proteins to consciousness, from fertilization to senescence, from the origins of the galaxy to the likely effects of the population explosion, from skin pigmentation to Piaget. This is not, however, an introduction to absolutely everything treated at a superficial level. Nor is it a textbook on human biology. There is, for example, no treatment of the special senses, nor even the structure of the nervous system as a whole. The author has a theme, the search for general principles of organization by which life is controlled and maintained. His material is selected to illustrate this theme and is discussed in considerable detail. Thus there is a fair amount on control systems, from single cells to the whole organism, including a useful review of the structure of DNA and its role in controlling protein synthesis. There are chapters on the physical and intellectual development of the individual and a section on the past and future growth of the human population. As one might expect, a large part of the book is concerned with the evolution of man, of his behaviour and of his culture.

Although specifically psychological topics form a relatively small proportion of the book, they are discussed in a stimulating and well informed manner. For example, the chapter on measurement of intelligence is nicely balanced, as good as anything of the same scope written by a professional psychologist and much better than chapters on the same topic in many introductory texts on psychology.

However, the main value of the book to psychologists (as indeed to all biologists), is in the background it provides to their studies. At a time when there is increasing concern over the specialization of degree courses, it provides an example of how a broad-based course may be presented without descending to the superficial. Given such an example, perhaps one might hope that courses of this kind may come to replace the hodge-podge of ill-assorted subjects inflicted on students at some institutions under the name of general degrees.

The book is beautifully produced on high quality paper. It has a profusion of well-chosen illustrations, bibliographies to each chapter, a glossary and subject and author indexes. In short, it is very good value indeed.

R. DAVIS

GRAY, J. A. *The Psychology of Fear and Stress*. London: Weidenfeld & Nicolson. 1971. Pp. 256. £1.75.

This book, which is published in the World University Library series, is primarily aimed at a wide non-specialist audience; the author, nevertheless, has both covered so broad a field that few specialists will be expert in all the topics covered and also produced provocative accounts of the literature within most of these topics. The book falls naturally into two halves, the first of which is concerned primarily with establishing the notion of fear, with its physiological correlates, and factors that affect fearfulness; the second, more psychologically oriented half, concerns the role played by fear in behaviour in general, and the place of fear in theoretical accounts of learning.

The first half of the book is written with admirable clarity, and covers a great deal of ground, including, for example, elements of genetics (as background to a discussion of the inheritance of fearfulness), and hormonal involvement in fear, stress, and sex. The author does not make any attempt to provide a comprehensive literature review, but rather follows an argument through in each chapter, providing evidence as the need arises. The result is an account which, although frequently controversial, is, more importantly, continuously readable.

Similarly, the later chapters of the book endeavour to make out a case for the author's own views, and will cause some reservations in specialists, although lay readers will once again find them both interesting and convincing. In summary, Gray wishes to argue that fear and frustration are closely linked, if not identical, responses, that they are mediated by the septo-hippocampal system, and that neurotic humans are essentially people with malfunctions of this system. The evidence used in developing this argument is occasionally not as convincing as it would seem to be from the text; for example, in discussing the effect of sodium amytal, Gray writes that the partial reinforcement extinction effect was, in an experiment of his own, virtually abolished in rats trained on amytal and extinguished on placebo. He then states that similar results were obtained by Ison and Pennes; but these authors in fact found no significant effect of the drug on the magnitude of the PREE, and finding (as did Gray) a large PREE in animals both trained and extinguished on amytal, concluded that "either the basic statements of frustration theory do not hold or the emotional reduction hypothesis of amobarbital action is incorrect". Thus, Ison and Pennes' results do not in fact provide the sort of support for Gray's views that one might be led to suppose. Again, in the succeeding chapter, on active avoidance, Gray takes the somewhat extreme view that the persistence of avoidance behaviour is to be explained by the secondary reward value of safety signals, and argues that such signals are rewarding "relatively independent of the degree of fear aroused in the situation". To support his position, Gray quotes a rather obscure experiment by Lawler, without referring to LoLordo's 1969 (*Psychological Bulletin*) review which (after specifically rejecting Lawler's experiment) concludes that there is as yet no convincing demonstration of positive conditioned reinforcement arising from aversive situations.

Gray's overview of the role of fear is not (at least not yet) a conventionally accepted one; the weaknesses in the presentation of his case are perhaps the inevitable result of his effort

to write a book that is both comprehensible by non-psychologists and of interest to his professional colleagues. In this effort he has been wholly successful; the book is stimulating and immensely readable, and should be read by everyone with an interest in psychology.

E. M. MacPhail

McFARLAND, D. J. *Feedback Mechanisms in Animal Behaviour.* London, New York: Academic Press. 1971. Pp. 279.

The application of control theory to biological systems is seen by some as being quite essential to their understanding, in something like the way that an ability to count is necessary for knowing whether one's change is correct. By others it is seen as yet one more fad whereby those with some faintly exotic expertise can rephrase what is known already in terms that serve to obscure it. Already there are several books which argue with varying degrees of cogency for the former point of view, but McFarland's is the first to deal specifically with the application of control theory to behaviour. The task he addresses in writing such a book is, implicitly at least, that of converting the doubters: equally clearly his success will be measured by whether he does so.

The first six chapters are a review of published applications of control theory to psychology. Yet despite a somewhat disingenuous claim in the preface that the author has tried to keep mathematics to a minimum, the book makes few concessions to the non-mathematical, is quite difficult, and is sometimes perfunctory in discussing the implications of the formal theory. Its main point therefore seems to be the presentation of McFarland's own experiments and theoretical treatment of animal motivation set in the context of some standard applications of systems theory to problems such as eye movements and various other aspects of motor control. There are also some slightly more off-beat asides such as the mention that detection of forgeries of works of art depends in part on whether the drawing of a line was under the smooth open loop (pre-programmed) motor control of the master, or of a forger trying to follow the original but being subject to the corrective movements inherent in feedback control.

Those people as yet unversed in control theory who are prepared to work at this book with the aid of supporting texts will be able to see how certain patterns of behaviour as a function of time derive from particular postulates as to the nature of the underlying control systems. Such patterns are not necessarily intuitively obvious without the formalism, but can be quantitatively described with it.

It is, however, the seventh and final chapter which is the most important. Entitled "Control theory and behaviour theory", it deals mainly with McFarland's most recent work and makes the case that we can only properly understand and predict motivated behaviour by formulating the problem in terms of a theory of multivariate control of the many potentially conflicting motivations by which animals are influenced. The argument is roughly that any one motivational system can be described by a differential equation whose terms can correspond formally to those of equations for mechanical or electrical systems. Such physical systems are usually discussed with reference to concepts of energy; thus a pendulum is said to oscillate by a reciprocating exchange between stores of kinetic and potential energy, with the characteristics of these stores appearing as terms in the equation. In just the same way terms in equations for motivational systems may be thought of as relating to stores of energy, and behaviour to the process of dissipating it. McFarland argues that the notion of energy has been used in one form or another by most workers who have discussed motivation (e.g. Freud, Hull, Lorenz) but that often the usage has led to contradictions because of confusions about whether this energy actually drives behaviour onwards and is used up or is merely a formal variable carrying no postulate as to its physical existence. For McFarland it is just such a variable, admissible only as an analogy to energy as used in physics. He goes on to postulate that although they say nothing about any embodiment of a mechanism in the brain, energy-like terms in his equations may allow prediction of some aspects of behaviour, and in particular of the results of interactions among motivational systems. The hypothesis is that an animal might in some way calculate a function which is a combination of suitably

weighted potential energy factors (corresponding to variables such as deprivation of food or water, internal sexual motivation, etc.) and kinetic energy considerations (related to the momentum of, and opportunity for, particular types of ongoing behaviour). The combined motivational system then is envisaged as optimizing this criterial function, and accordingly McFarland proposes that it can be analysed in terms of a theory of optimal control capable of producing many different outputs via a single behavioural common path.

It remains to be seen whether this idea can be translated into a working theory. Empirical determination of any optimality criteria, or of ways in which such criteria could be specified or met, has yet to be carried out. Indeed, the question as to whether this is the best, or even a useful, approach to understanding how an animal chooses what to do is still in doubt. But there is no question that this is a bold enterprise which may illuminate the problems of displacement, frustration and behavioural choice that remain a puzzle for students of animal motivation.

KEITH OATLEY

DARTNALL, H. J. A. (Ed.). *Photochemistry of Vision*. Berlin: Springer-Verlag. 1972. Pp. 810. $59.30.

Even in these days of rapid scientific advance, and looseleaf textbooks, this, "Dartnall's" volume of the enormously long Handbook of Sensory Physiology is bound to become a classic. In earlier textbooks to which he has contributed, Dartnall has adopted the somewhat unusual practice of including much unpublished material, and here his 19 contributors have to a large extent followed his example. Perhaps the most shining example is that of Yoshizawa, whose work on low-temperature spectroscopy of visual pigments has hitherto been almost totally unavailable: and in a work of this length, Yoshizawa has been able to include many interesting technical experimental details, which do not often escape the censorship of journals' editors. The aim of the work is to be exhaustive, and the contributors have gone to considerable length to ensure that their results should be easily intelligible. Thus, Abrahamson and Japar preface the volume with an introduction to the quantum-mechanical principles of interactions of light and matter, starting from the simplest considerations and leading on to the description of the molecular wave functions (with diagrams for those not gifted with mathematical intuition): and though, in 25 pages, it is obviously impossible to do justice to this branch of physics, the uninitiated biologist will find that the introduction enables him to make considerably more sense of the energy diagrams associated with the light absorption of retinal pigments than has previously been the case. "Classic" photochemistry of rhodopsin is adequately summarized by Morton, Abrahamson and Weisenfeld, and by Dartnall; Shaw contributes a special chapter on circular dichroism and optical rotatory dispersion. These chapters only occupy 200 out of nearly 750 pages. Though much of what they contain is highly specialized, the investigations form the basis of our understanding of the visual system as a whole, and should be mandatory reading for any student whose interest is connected with sight. The studies are particularly relevant to problems of excitation of photo receptors, and the implications of the bleaching of rhodopsin, the associated changes in thermodynamic quantities, the Cotton effect and so on, are carefully considered by the various authors. Some of the suggestions in these sections are already out of date: one wishes that recent data on X-ray diffraction and rapid photo-induced changes in rhodopsin dichroism *in situ* could have been included; but whatever the final deadline for manuscripts the production of a volume such as this is a fairly lengthy process, and speculations are bound to date; this does not detract from the work as a whole.

Falk and Fatt contribute a chapter on the link between light absorption and neural excitation, in an impressive article which summarizes the known facts about the way in which absorbed light energy is dissipated in receptors, and their own work on impedance changes of outer limbs. These reveal rapid increases in intracellular conductance, possibly of disc membranes, that may be circumstantially related to an early stage in the amplification mechanism which leads to the neurophysiologists' photo-responses. Falk and Fatt stop before the recent work by Bownds, Hagins and Yoshikami, and Bitensky *et al.* which suggests

that calcium fluxes or changes in cyclic AMP may be associated with such early stages of excitation: this obviously has important implications for their findings. A provocative and thoughtful section is included on the difficulties of relating present knowledge about photoreceptor currents and the mechanisms of synaptic transmission to the known reliability and sensitivity of the visual system.

The topics mentioned above occupy only one third of the volume. The remainder is devoted to articles showing how the study of photochemistry has illuminated fields in zoology and ecology. Crescitelli traces the development of visual cells and visual pigment in a variety of vertebrates, and shows how these are related to the evolutionary history of the species, and in the case of fish to the marine and freshwater environment. The same theme is treated in more detail by Bridges, who describes the rhodopsin-porphyropsin system with a depth of understanding, and considers not only the environmental factors which cause this interconversion in Salmonidae but also the retinal mechanisms which enable the transformation to take place. Baumann reviews work on regeneration of rhodopsin, including the cycle between rods and pigment epithelium, and in the continuing growth of rods.

Work on cone pigments is described in detail by Liebman, who gives the theoretical basis for microspectrophotometry, and also includes much previously unpublished work, while Rushton surveys the principles of reflection densitometry as applied to man, and reviews the correlations established by these means between pigment concentrations and visual functions such as dark adaptation and colour vision. Nor are invertebrates neglected: Eakins and Goldsmith review the structure and the pigments of a variety of invertebrate forms. Goldsmith quotes Hartline, "If it's simple, it's not an eye", but in spite of this, advances in our comprehension seem certain to stem from further study of these very diverse light-sensing organs. Finally, stable pigments, and those not contained in photo-receptors, are reviewed by Muntz and by Hara and Hara, who give a complete account of cephalopod retinochromes. It can be seen that the subject of the volume is treated very catholically, and perhaps nowhere more so than in Lythgoe's first chapter, which deals almost entirely with the distribution of light under various bodies of water, and its possible effect on vision: but if this ecological excursion is straying a bit from the theme, the author makes amends with his second chapter, a list of all known visual pigments.

The presentation of the material is exceptional, the cross references between chapters frequent, and the English plain and straightforward, greatly adding to the ease of reading the book. For this, as various footnotes testify, the Editor must have taken a great deal of work upon himself. Altogether, this is a volume of such unusually high quality as to fully justify the effort, and it will no doubt remain a standard for many years to come.

G. B. Arden

Victor H. Denenberg (Ed.). *Readings in the Development of Behaviour*. 1972. Stamford, Conn.: Sinauer Associates Inc. Pp. 483 + ix. £2.90.

This is a collection of 65 papers, with long introductory passages from the editor. The papers are not abridged. Many are short; there are sixteen from *Science* alone. The *Journal of Comparative and Physiological Psychology* is also well represented with 11 papers; others come from a variety of medical journals (*Pediatrics, Human Development*, etc.).

The coverage of clinical studies, as well as experimental work on animals, is an important feature of the collection. There are 13 "chapters" in all, each dealing with a particular area of developmental research (critical periods, prenatal effects, environmental enrichment, etc.): the section on critical periods contains a study of the intelligence of children with diabetes of early and late onset, that on infantile stimulation a cross-cultural study of the effects of stress in infancy on adult stature, and an experimental study on the consequences of handling in infancy on the development of premature babies. This carefully contrived, juxtaposition of clinical and comparative studies demonstrates genuinely important parallels, and is a most satisfying feature of the collection.

Many of the papers are essential reading in the area they represent: one cannot understand infantile stimulation without having read Levine's work on the effects of handling on adult

corticosterone output, nor can one understand sexual differentiation without working through W. C. Young's classic papers, or the development of perception without coming across Fantz and Held and Hein. There are occasional choices I regret; it is a pity to omit the simple and central study of Resko on plasma testosterone levels in the newborn rat when room is found for a report on the suppression of mating in the female by perinatal oestrogen, an unphysiological effect and one that can only confuse matters unnecessarily for undergraduates. Nor would I wish to have any more publicity given to the editor's study of the effects of infantile stimulation on the survival of rats placed on a "terminal deprivation schedule".

As regards Denenberg's introductions and bridging passages, much of what he says is valuable commentary on the papers, linking them together and bringing out the general nature of some of the problems of the area. His writing suffers from a tendency to lapse into jargon of an unilluminating kind; e.g. he describes "learning" on different pages, and sometimes the same page, as a phenomenon, a mechanism, a process variable, a psychological variable and a psychological behaviour. His discussion of research on sexual differentiation lacks precision and accuracy. He begins on page 11 by saying that "the presence or absence of certain sex hormones (androgen for the male and estrogen for the female) can determine whether an organism will behave in a male-like or a female-like fashion". This is misleading: the presence of androgen suffices to produce a male; its absence a female. Oestrogen has no known part in the differentiation of the female. The main discussion of the area (p. 113) presents an over-simplified account couched in terms of the "male brain" and the "female brain". He ignores the fact that full masculine and feminine sexual responsiveness can co-exist in the same animal (in, for example, male rats treated neonatally with androstenedione), a fact which makes nonsense of a general distinction between the "male" and "female" brain. He assumes that all behavioural changes produced by manipulations of testosterone in infancy must be due to differentiation in the central nervous system, simply because this is true of the differentiation of sexual cyclicity, and states that the "feminized" male shows a reduced tendency to mount. It does not. The components of the mating pattern that decrease in animals castrated on the day of birth are intromission and ejaculation, and the fact that these, but not mounting, decrease is one of several lines of evidence suggesting that it is peripheral, not central, differentiation that produces the changes in male sexual behaviour found in the infant castrate.

Denenberg pleads for more "two-way conversations" between workers in different areas but his comments on psychogeneticists (p. 48) and psychiatrists and psychoanalysts (p.399) are not likely to help such communication. We learn that psychogeneticists are naive, out-of-date and badly educated, though without ever discovering exactly why the editor thinks this, or who exactly he is talking about. In his discussion of maternal influences he criticizes "the psychiatric, especially psychoanalytic, bias of a number of older investigators, which led them to assume the importance of the mother ... rather than attempting to test the validity of the assumption". There are heads this cap might fit, but it is hardly fair to single out Bowlby as an example; his 1956 study of the after-effects of confinement to a tuberculosis sanatorium in the first years of life is one of the most carefully planned and rigorous clinical experiments in the area, and contains in the discussion the remark that "statements implying that children who are brought up in institutions or who suffer serious privation or deprivation in early life commonly develop psychopathic or affectionless characters (e.g. Bowlby, 1944) are seen to be mistaken". Nor would one guess from Denenberg's remarks that Bowlby has been successfully encouraging for years the comparative "biobehavioural" approach that Denenberg here espouses. The other psychiatrist singled out is Spitz, who is criticized because "since Spitz refused to reveal the locations of the children and would not allow others to examine his raw data, his reports cannot be considered to be part of the body of scientific knowledge—which demands independent investigation and verification . . ." It is the elementary duty of every clinical investigator to safeguard the privacy of his subjects, and the comment about independent investigation is misplaced: it is independent tests of the same generalization that are essential for the march of science, not independent checks of every piece of raw data.

But this is essentially a book of readings, and the readings themselves are usually comprehensive and well organized. There are not many undergraduates intent on a career in clinical psychology who think of this type of "biobehavioural" research as a possible area of specialization. Denenberg's collection could well add to their number, and thereby serve a very useful purpose.

R. F. DREWETT

MYERS, R. C. (Ed.). *Methods in Psychobiology: Laboratory Techniques in Neuropsychology and Neurobiology.* London: Academic Press. 1971. Pp. xi + 356.

"In fact the rapid developments in biochemistry, neurophysiology and neuropharmacology have already led to a situation whereby the psychologist is forced either to restructure his theoretical and practical thinking or to lapse into an otherwise traditional rather than modern approach to experimental investigation." This chill warning from the preface underrates advances still being made in traditional fields with traditional methods—but at the same time one cannot doubt there are many senescent psychologists (i.e. psychologists already with Ph.D.s) to whom this "cookbook" of techniques, procedures, recipes and principles is going to bring timely relief. Eleven chapters by 12 established artificers in the behavioural sciences spell out carefully the needs and procedures for the opening assault on the New Psychology: where to write for the equipment (in appendices to most chapters), how to handle experimental animals and manage them under anaesthetic (Myers), how and where to implant stimulating electrodes (Pellegrino and Cushman), how to make lesions in the cerebral cortex (P. M. and D. R. Myer) or beneath it (Thompson), how to record the brain's electrical activity (Cooper), how to stimulate (Ervin and Kenney) or inject it (Myers again) and how to work out afterwards just where one went astray (Wolf: Elementary Histology for Neuropsychologists). The remaining two chapters (Wayner, Falk) are largely admonitions on scientific method presumably aimed at an undergraduate audience, but I suspect their paternal and simplistic homilies will not impress a generation reared on Kuhn (and worse).

Many readers would undoubtedly have wished for more guidance on simple biochemical and pharmacological procedures; and there is nothing on the use of isotopes, or on the measurement of blood pressure, heart rate or body temperature, nor on the measurement of motor activity—the latter being by all accounts a particularly chaotic pursuit in psycho-pharmacology at the present time. These gaps, however, will very possibly be remedied by further volumes which the publishers seem to have in mind, and of more concern may be lacunae within topics already dealt with: Robert Thompson, for example, seems to have had the misfortune to have gone to press too soon to include the Gold-Sclafani knife-cut technique—undoubtedly the most important lesioning innovation of the decade. And of course anyone already in the game will know of labour-saving ploys which the contributors have overlooked; this reviewer, for example, has long ago discovered that the simplest way to make bipolar electrodes is not by cementing together two straightened lengths of wire (p. 219); one need merely tie two strands to the doorknob and spin them in the Black and Decker for two seconds. And hand-held microinjections do not really require exceedingly careful pressure on the plunger (p. 236)—surely micrometer syringes can be got in Indiana! Nor in this age of ammonium fixatives need one wait a full 30 minutes (p. 149) before rinsing bromide prints—30 seconds would be closer to the mark. I could go on, but the last thing I should want to do is belittle the good sense and quite exceptional value of this collection—for the skilled practitioner no less than for the beginner. Here at last is a commodity that might truthfully promise to repay its purchase price many times over. And it is for just this reason that one must hope prospective readers will not be put off by its clumsy jargon, malapropisms and innumerable printing errors, these coming as something of a surprise from a publisher elsewhere associated with the very highest standards of editing—as every reader of this *Journal* must surely know. But be forewarned: at least two contributors speak approvingly of the "facile" methods developed by a colleague; and "virile diseases" will be found to mean viral infections, not something you overlooked in Krafft-Ebing.

L. J. HERBERG

Publications Received

BEIDLER, L. M. (Ed.). *Chemical Senses 1. Olfaction.* Handbook of Sensory Physiology. Vol. IV. Berlin: Springer-Verlag. 1972. Pp. 518. $45.70.

BOWER, G. H. (Ed.). *The Psychology of Learning and Motivation.* Vol. 5. New York: Academic Press. 1972. Pp. 388. $14.50.

DAWES, R. M. *Fundamentals of Attitude Measurement.* New York: John Wiley. 1972. Pp. 165. £3.50.

FORBES, T. W. (Ed.). *Human Factors in Highway Traffic Safety Research.* Chichester: John Wiley. 1972. Pp. 419. £7.85.

FREUD, A. *et al.* (Eds). *The Psychoanalytic Study of the Child.* Vol. 26. London: The Hogarth Press. 1972. Pp. 616. £5.00.

GELDARD, F. A. *The Human Senses.* 2nd ed. New York: John Wiley & Sons. 1972. Pp. 584. £5.25.

LINDSAY, P. H. AND NORMAN, D. A. *Human Information Processing.* New York: Academic Press. 1972. Pp. 737. $11.95.

LOWE, G. R. *The Growth of Personality: from Infancy to Old Age.* Harmondsworth: Penguin Books. 1972. Pp. 272. £0·45.

LORENZ, K. *Studies in Animal and Human Behaviour.* Vol. II. London: Methuen. 1971. Pp. 366. £4.25.

MASTERSON, J. F. *Treatment of the Borderline Adolescent.* New York: John Wiley. 1972. Pp. 289. £5.50.

MYERS, J. L. *Fundamentals of Experimental Design.* 2nd ed. Boston: Allyn and Bacon Inc. 1972. Pp. 465.

RUTTER, M. *Maternal Deprivation Reassessed.* Harmondsworth: Penguin Books. 1972. Pp. 175. £0·50.

SCHACHTER, S. *Emotion, Obesity, and Crime.* New York: Academic Press. 1971. Pp. 195. £3.50.

STELLAR, E. AND SPRAGUE, J. M. (Eds). *Progress in Physiological Psychology.* Vol. 4. New York: Academic Press. 1971. Pp. 363. £7.00.

TOBIAS, J. V. (Ed.). *Foundations of Modern Auditory Theory.* New York: Academic Press. 1972. Pp. 508. $24.00.

TOCH, H. *Violent Men: an Inquiry into the Psychology of Violence.* Harmondsworth: Penguin Books. 1969. Pp. 304. £0·40.

WASON, P. C. AND JOHNSON-LAIRD, P. N. *Psychology of Reasoning: Structure and Content.* London: B. T. Batsford. 1972. Pp. 264. £3.50.

WINOGRAD, T. *Understanding Natural Language.* New York: Academic Press. 1972. Pp. 191. $8.95.

Author Index

559

Notes for Authors

The Journal publishes original papers on experimental work in all branches of human and animal psychology. Papers dealing with the mechanisms underlying behaviour are particularly welcome and theoretical papers on such topics are acceptable provided they relate to experimental findings. Short notes on apparatus likely to be of interest to experimental psychologists are also acceptable, provided that the design of the apparatus is an original contribution by the author.

Papers should be submitted in duplicate, one copy to be a *typewritten original*, preferably on A4 paper (219 × 297 mm), in double spacing with wide margins.

Title The title should be as concise as possible and should appear as a separate sheet together with the name(s) of the author(s) and the full postal address of his institution. A short running title of not more than 40 characters (including spaces) should be indicated if the full title is longer than this.

Abstract An abstract of 50–200 words should follow the title page as a separate sheet.

Headings and sub-headings for different sections of the paper should be clearly indicated. Authors should follow the conventions used in current issues of the Journal.

Acknowledgements should be as brief as possible and should appear as a separate sheet at the end of the text.

Footnotes should be avoided unless absolutely necessary. Essential footnotes should be indicated by superscript figures in the text and collected as a single page at the end of the manuscript.

References in the text should be indicated with the year of publication, in parentheses after the author's name. If more than one article by the same author in the same year is quoted, *a*, *b*, etc., should follow the year. When a paper has been published jointly by two authors, both names should be given in the text. For more than two authors, quote all names at the first mention and use the first author '*et al.*' for subsequent mentions. References to unpublished work should be avoided since they are of little use to readers. A list of the references quoted in the text should be given at the end of the paper in alphabetical order of author's names, commencing as a new sheet. Titles of Journals and books should be given *in full* thus:

WOODWORTH, R. S. (1901). On the voluntary control of the force of movement. *Psychological Review*, **8**, 350–9.

GALTON, F. (1893). *Inquiries into Human Faculty and its Development*. London: Macmillan.

GARRETT, M. and FODOR, J. A. (1968). Psychological theories and linguistic construction. In DIXON, T. R. and HORTON, D. L. (Eds), *Verbal Behavior and General Behavior Theory*. Pp. 451–77. Engelwood-Cliffs, N.J.: Prentice-Hall.

Tables should be kept to the minimum consistent with clarity of exposition. Each table should be typed double spaced as a separate sheet, giving the heading, e.g. TABLE II, in Roman numerals, followed by the legend, followed by the Table. Make sure that appropriate units are given.

Instructions for placing the Table should be given in parentheses in the text, e.g. (TABLE II about here).

i

Figures are expensive to reproduce and should only be used when essential. The same data should not be presented both as a figure and in a table. Where possible, related diagrams should be grouped together to form a single figure. Figures should be drawn to professional standards in Indian Ink on Bristol Board of A4 size, and it is recommended that the linear dimensions of figures be approximately twice those intended for the final printed version. Lettering to appear in the figure should be done on tracing paper placed over the figure. Make sure that axes of graphs are properly labelled and that appropriate units are given.

The legend should be typed on a separate sheet of paper headed, e.g. FIGURE 2, in Arabic numerals. Instructions for placing the figure should be given in parentheses in the text, e.g. (FIGURE 2 about here).

Statistics Greek letters should be used to represent population parameters and Roman letters for sample values. Results of statistical tests should be given in the following form:

$$F = 25.3, \, df = 1,9, \, P < 0.001.$$

Similarly for t, χ^2 and other tests.

Units The Système Internationale (SI) will be used for all units. (If measures such as inches are used, then SI equivalents should be given.) If in doubt, refer to *Units, Symbols and Abbreviations*, published by the Royal Society of Medicine, 1 Wimpole Street, London, W.1.

The author will normally receive a decision on his paper within 3 months of receipt and if it is accepted it will normally be published 3 to 6 months later. The date of the receipt of the manuscript will be printed. Where minor revision of a paper is requested the original date of receipt will appear, provided that a satisfactory revision is received within 1 month of the request.

Contributors receive 25 copies of their paper free. Additional copies may be ordered on a form provided by the publishers at the time proofs are sent to authors.

The Nature of Reinforcement

Edited by ROBERT GLASER

Professor of Psychology and Education
Co-Director, Learning Research and Development Center
University of Pittsburgh
Pittsburgh, Pennsylvania

This book presents a comprehensive review of theory, research, and application in the field of reinforcement. The topic is particularly relevant today, not only because of the nature of current theoretical and experimental activities, but also because of the increasingly wide use of the principles of reinforcement in practical affairs.

The volume contains contributions from many of the most outstanding workers in the field, including:

- W. K. Estes, who examines theoretical issues in the interpretation of reward in human learning;

- Frank A. Logan, who discusses reinforcement in the context of incentive theory;

- Richard C. Atkinson and Thomas D. Wickens, who present a theory of the role of reinforcement in terms of information processing concepts and in the context of a general theory of memory;

- David Premack, who extends his formulation of relative response probabilities to an analysis of punishment;

- Albert Bandura, who examines the functional properties of vicarious reinforcement and self-reinforcing events.

1971, 379 pp., £6.05

Academic Press

NEW YORK AND LONDON
111 Fifth Avenue, New York, NY 10003
24-28 Oval Road, London NW1

OCCUPATIONAL PSYCHOLOGY

Editor: ISABEL BLAIN

1972   Volume 46 No. 2

Contents

The British Journal of Psychology

Volume 63 Part 3   August 1972

CONTENTS

Primate Behavior

Developments in Field and Laboratory Research
Volume 2

Edited by LEONARD A. ROSENBLUM

*Director, Primate Behavior Laboratory, Department of Psychiatry,
Downstate Medical Center, Brooklyn, New York*

This is the first regular publication in the United States devoted to the burgeoning field of primate behavior. The series contains reviews of the literature and extended research programs, conceptual and theoretical papers, and major individual research efforts. Its scope ranges from considerations of particular dimensions of primate behavior (e.g., learning) subjected to detailed quantitative analysis within the rigorous confines of the laboratory to the broad sweep of the behavior of a species observed under ecologically complex field conditions.

This serial publication will be valuable to a wide range of researchers and scholars concerned with the behavior of primates—psychologists, anthropologists, zoologists, psychiatrists, physiologists, pharmacologists, veterinarians, and space technologists.

1971, 280 pp., £6.30

Number 3

Categorization and Social Judgement

J. R. Eiser
Department of Psychology
University of Bristol, England
and Wolfgang Stroebe
Department of Psychology
University of
Massachusetts, Mass., U.S.A.

Winter 1972
approx. 250pp., £3.50

In this book the authors propose a new theory of social judgement. Although based on the principles of psychophysical judgement as are most other social judgement theories, this new approach stresses two important factors; that the stimuli used in social judgement experiments typically differ in terms of more than one attribute, and that the rating dimensions for these stimuli are often highly value-laden.

This theory is proposed only after an exhaustive examination of related data. The first part of the book reviews the effects of context on judgements of physical stimuli with a brief examination of some theories to account for these effects. The authors then study the work on the effects of value on judgements of physical magnitude before reviewing the major theories of social judgement and the research they have stimulated. In the light of the conclusion that none of these theories can successfully account for all research findings in this area the book then presents its own original theoretical approach.

Contents: Theories of absolute judgement in psychophysics.
Stimulus distinctiveness and stimulus relevance.
The effects of incidental stimulus variation on absolute judgements.
Theories of social judgement.
An accentuation theory approach to social judgement.
Social judgement and phenomenal causality.
References.
Subject index.

CONTENTS